# DNA Replication

Instructor: Maureen Hoatlin
Biochemistry and Molecular Biology
Email: hoatlinm@OHSU.edu

# Lecture Overview

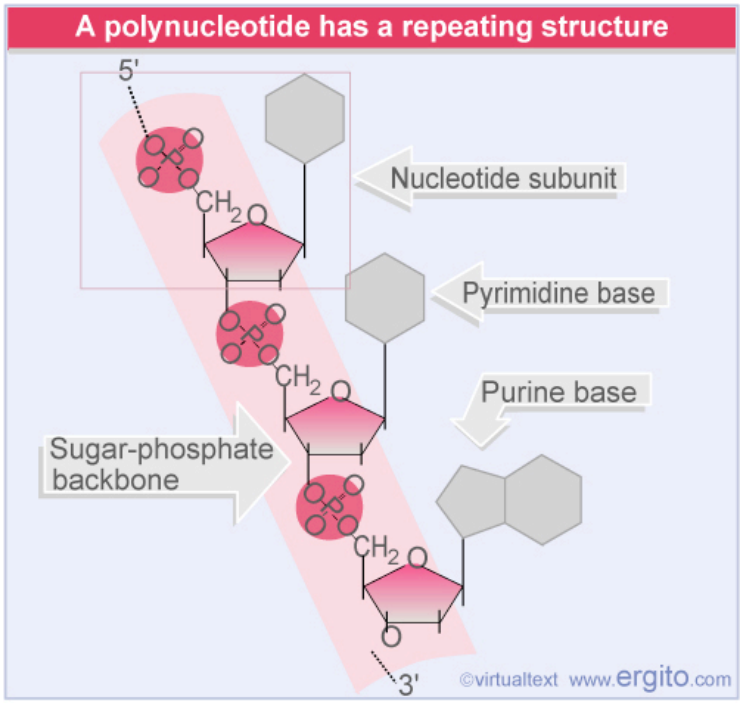
- DNA Review
- Essential components of the replisome
  - Genetic analysis identifies components
- Polymerases, SSB proteins, clamps and clamp loaders, helicases
- How DNA replication proceeds
- How DNA replication starts
- What is known and what remains to be discovered

# Significance of DNA Replication

- At least 38 diseases are caused by defects in DNA replication, 40 by mutations in genes required for DNA replication or repair
- Many drugs used to treat diseases caused by viruses are targeted to DNA replication
- Many chemotherapy agents are targeted to DNA replication
- Deconstructing DNA replication is central to efforts aimed at developing new diagnostic tools and new treatments for cancer

# General Structure of DNA

- Long strands of polymerized nucleotides
- Base, sugar and a phosphate

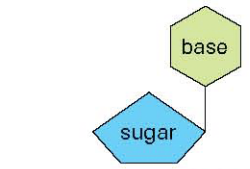



NOMENCLATURE The names can be confusing, but the abbreviations are clear.

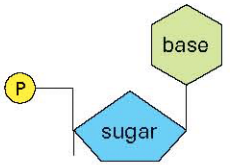
| BASE | NUCLEOSIDE | ABBR. |
|---|---|---|
| adenine | adenosine | A |
| guanine | guanosine | G |
| cytosine | cytidine | C |
| uracil | uridine | U |
| thymine | thymidine | T |

Nucleotides are abbreviated by three capital letters. Some examples follow:

AMP = adenosine monophosphate
dAMP = deoxyadenosine monophosphate
UDP = uridine diphosphate
ATP = adenosine triphosphate




BASE + SUGAR = NUCLEOSIDE



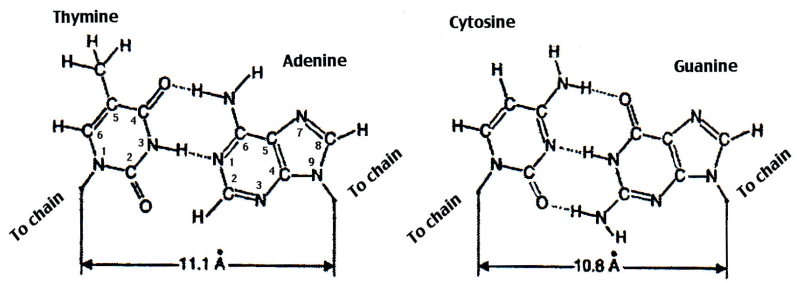

BASE + SUGAR + PHOSPHATE = NUCLEOTIDE

From Alberts text

2009

# DNA Bases are Paired

- A pairs with T, C pairs with G
- Purine/pyrimidine pairing results in a consistent overall dimension of the DNA duplex
- Base pairing is stabilized by hydrogen bonds; G:C rich helices are more stable than A:T rich ones
- Helix dimension has implications for DNA replication fidelity and for detection of certain types of DNA damage

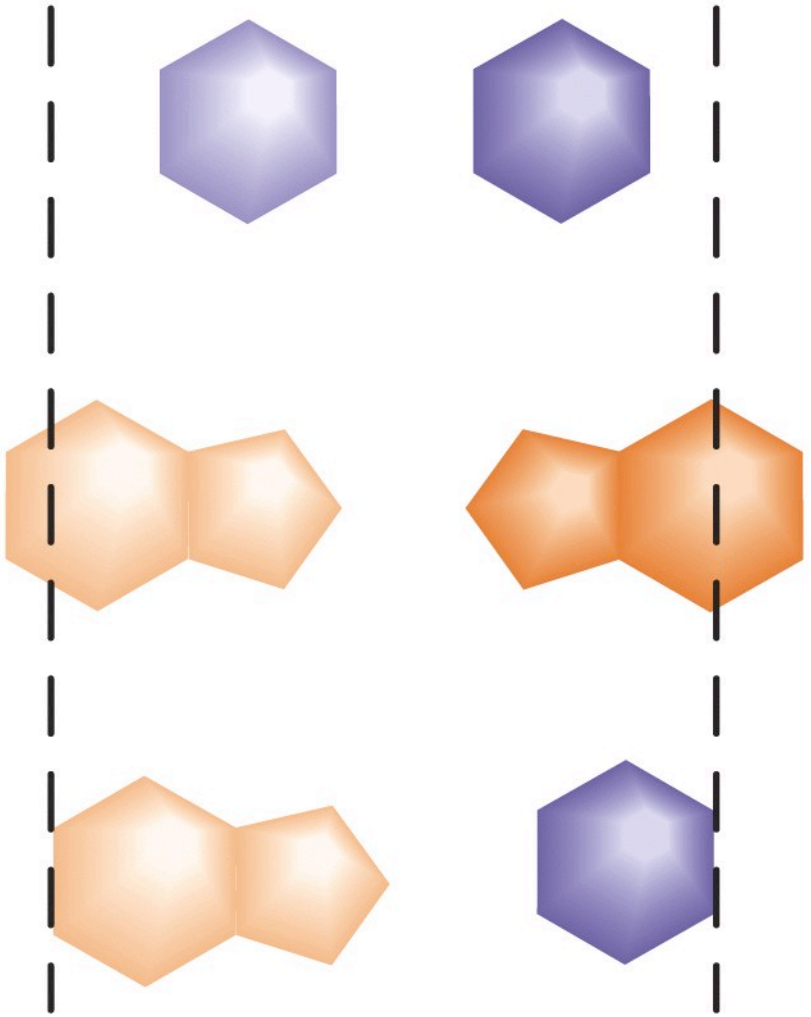



2 H bonds  3 H bonds

**Pyrimidine + pyrimidine: DNA too thin**

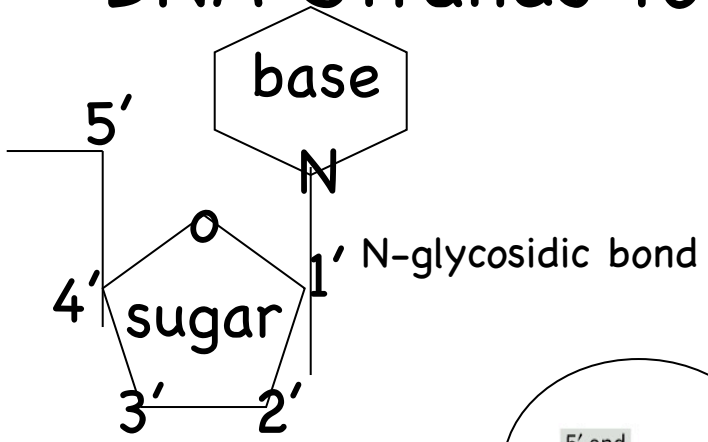
**Purine + purine: DNA too thick**

**Purine + pyrimidine: thickness compatible with X-ray data**



Hoatlin Fall 2009

# DNA Strands form an anti-parallel helix



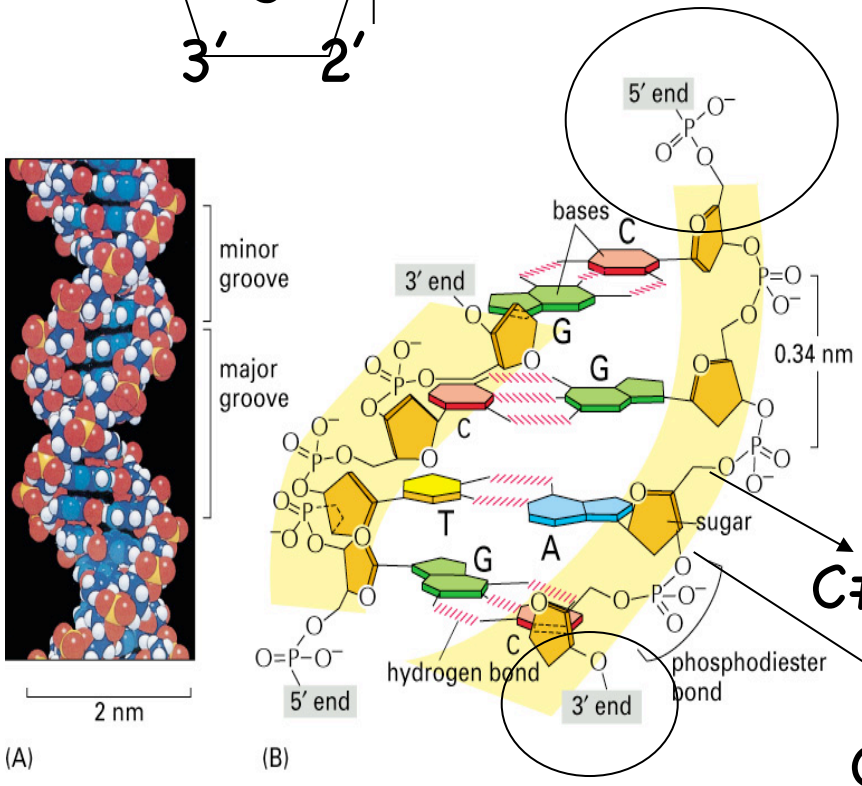




Figure 4–5. Molecular Biology of the Cell, 4th Edition.

Two DNA Strands are twisted together in a helix, called a double helix

Sugar phosphates are on outside of helix, bases on the inside

A bulky two-ring base (purine; A&G) is always paired with a single-ring base (pyrimidine; T &C). Heterocyclic rings are flat.

Helix is stabilized by hydrogen bonding between the bases and stacking forces

Nucleotides are linked by covalent phosphodiester bonds (3′-OH to 5′-PO4)

**DNA ends have chemically-defined polarity**

Note: review chemical bonds of helix

# Base-Pairing Underlies DNA Replication and Repair

Each strand can act as a template for duplication- nucleotide A pairs with T, and G with C

Each template nucleotide is recognized by a free complementary nucleotide

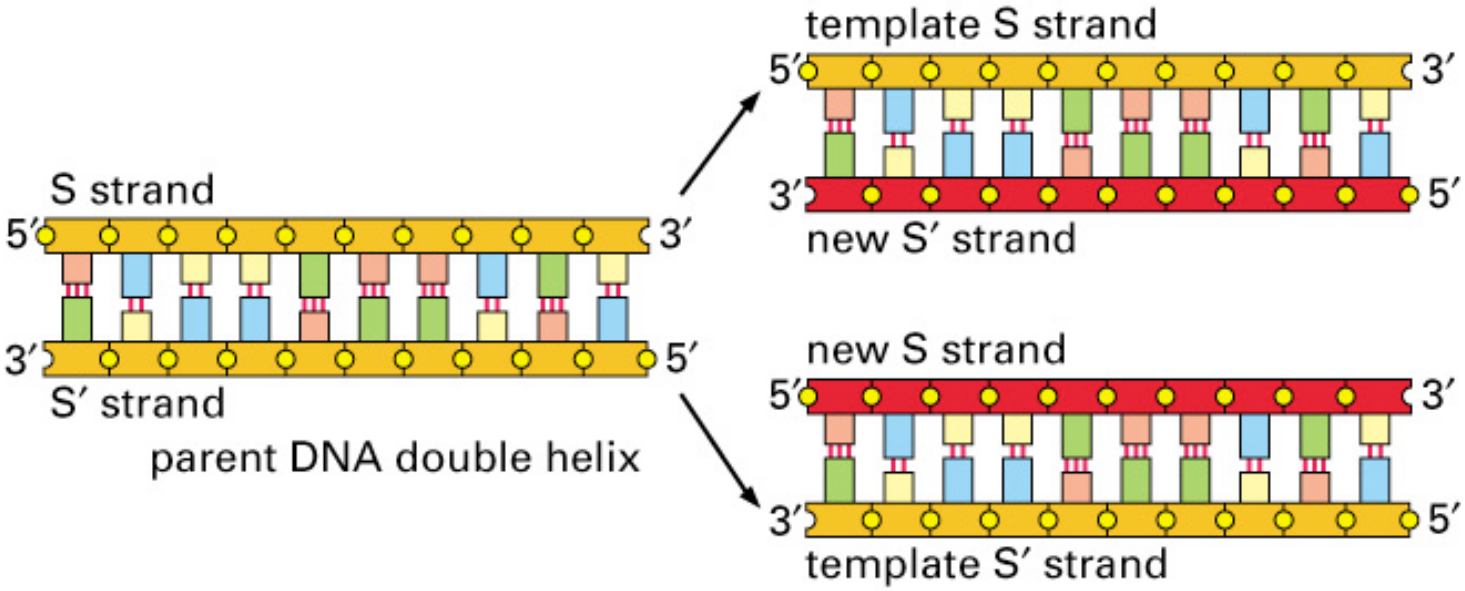



Figure 5–2. Molecular Biology of the Cell, 4th Edition.

# Overview of Replication

- Initiation (covered in lecture #2)
  - Recognition of an *origin* by a complex of proteins. Parental strand separation and stabilization. Synthesis can start at the *replication fork*.
- Elongation (covered in lecture #1)
  - the parental strands unwind, the *replisome* (complex of proteins) moves along DNA, daughter strands synthesized
- Termination (covered in lecture #2)
  - Termination reactions preceed separation of duplicated chromosomes

Our Attention is here

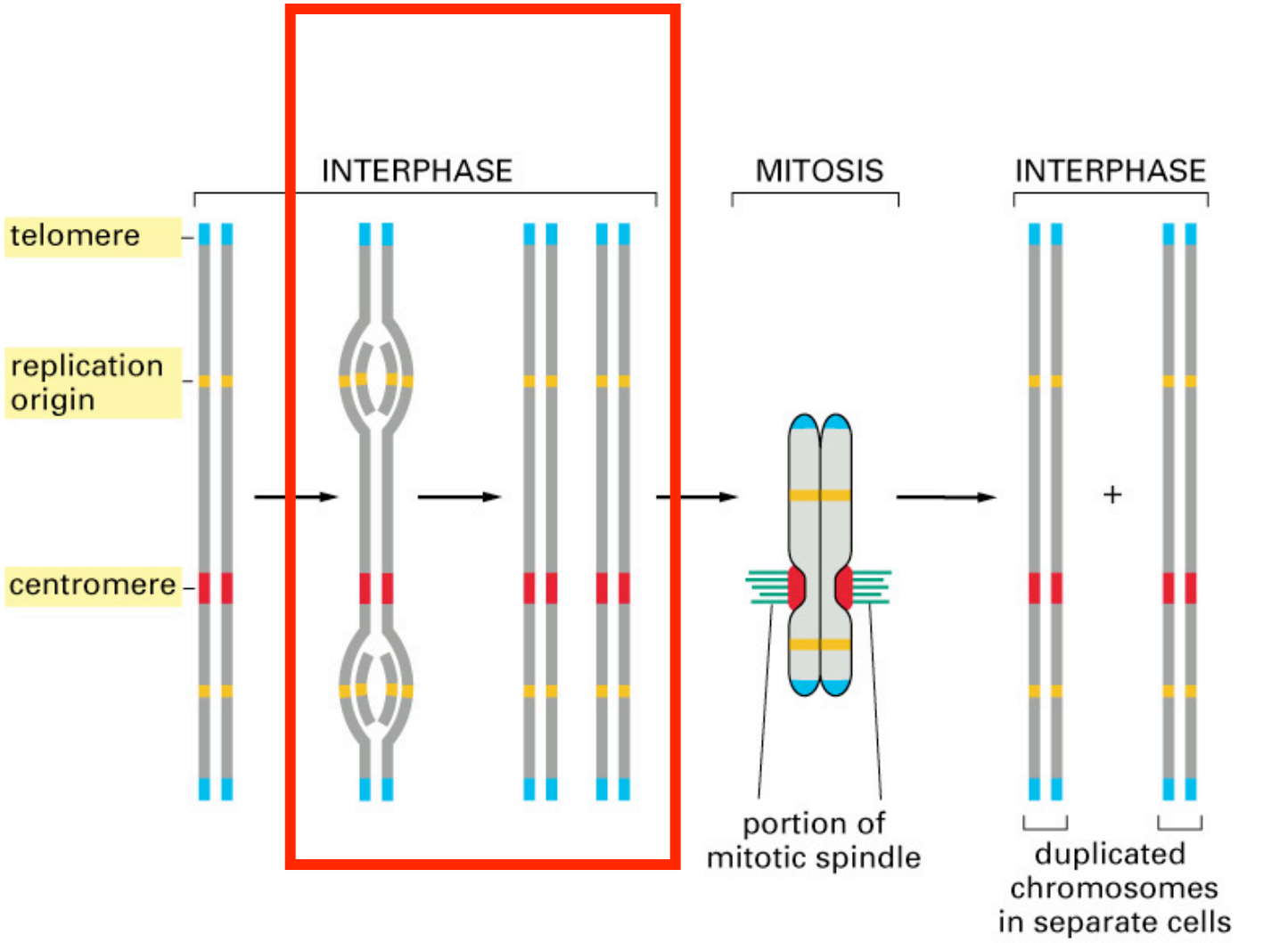



Figure 4–22. Molecular Biology of the Cell, 4th Edition.

# Mutation Rates Are Very Low

Bacteria-estimates from rapidly growing cultures are 1 nucleotide change in $10^9$ per cell generation

Mammals-similar rates per round of replication. Ex: human genome ($3 \times 10^9$) could have 3 mutations/cell division

Some nucleotide changes are silent, while others can have major consequences for the species or the individual.

Maintaining genomic integrity is dependent on accuracy of duplication and distribution of DNA to daughter cells, and on response to and repair of DNA damage.

# Components of Cellular Replicases

| *Replicase Component* | *E. coli* | *Eukaryotes* | *Archaea* | *T4 Phage* |
|---|---|---|---|---|
| **Polymerase** | Core (αεθ) | Pol δ/Pol ε | Pol δ | gp 43 |
| **Sliding Clamp** | β clamp | PCNA | PCNA | gp 45 |
| **Clamp Loader** | γ complex | RFC | RFC | gp 45/62 |

…variations on a theme

*E. coli* is best understood

# Dissecting Replication

- dna mutants: large set of *E. coli* mutants that distinguish two stages of replication
- Inability to replicate is lethal. Mutants in replication are obtained as *conditional lethals*. These mutants are able to replicate under permissive conditions (e.g., 25°) but are defective under non-permissive conditions (e.g., 39°).

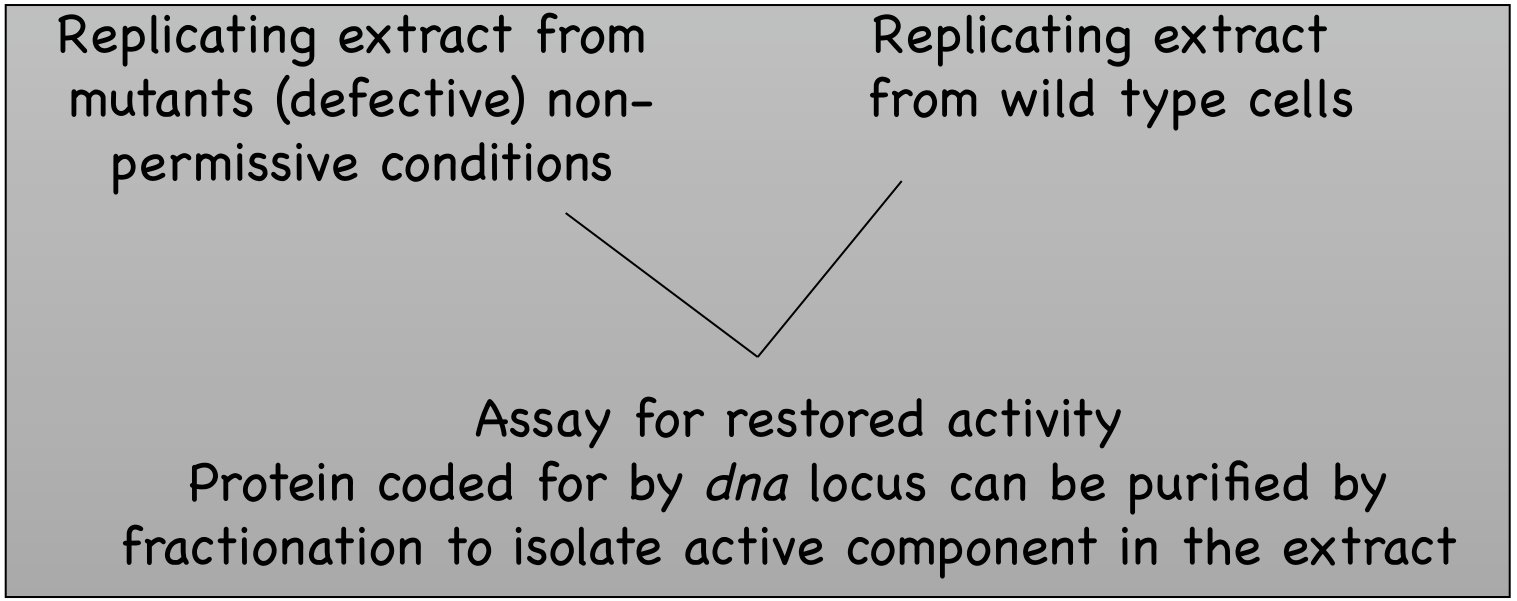
# dna Mutants

- Quick stop mutants (majority)
  - Cease replication immediately under non-permissive temperatures
  - Typically defective in components of the replication apparatus (enzymes needed for elongation, or essential precursors)
- Slow stop mutants (minority)
  - Complete current round of replication but cannot start another
  - Defective in events leading to initiation of replication at the origin

Examples to follow

# Identifying the Components of the Replication Apparatus

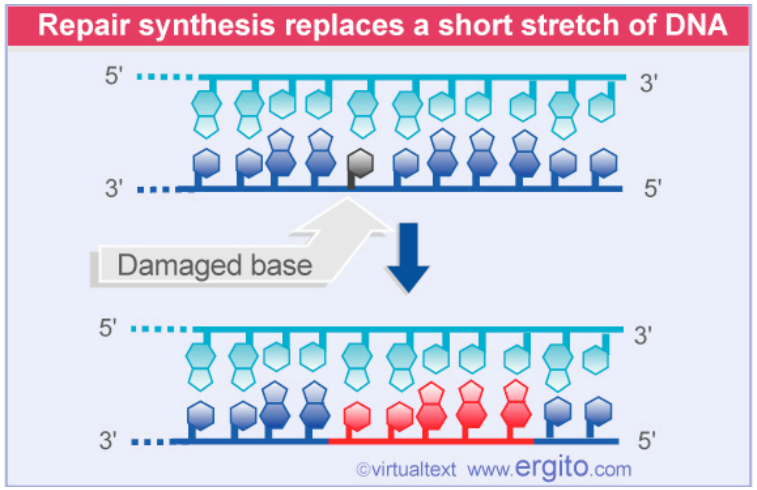
- *In vitro* complementation




- Bacterial replication proteins purified, eukaryotic replication systems analogous but incomplete and unsatisfying.

# DNA Polymerases make DNA

- Two basic types of DNA synthesis
  - Semiconservative replication
  - DNA synthesis in a DNA repair reaction




Damaged base is excised and new DNA is synthesized

# DNA Polymerases Control the Fidelity of Replication

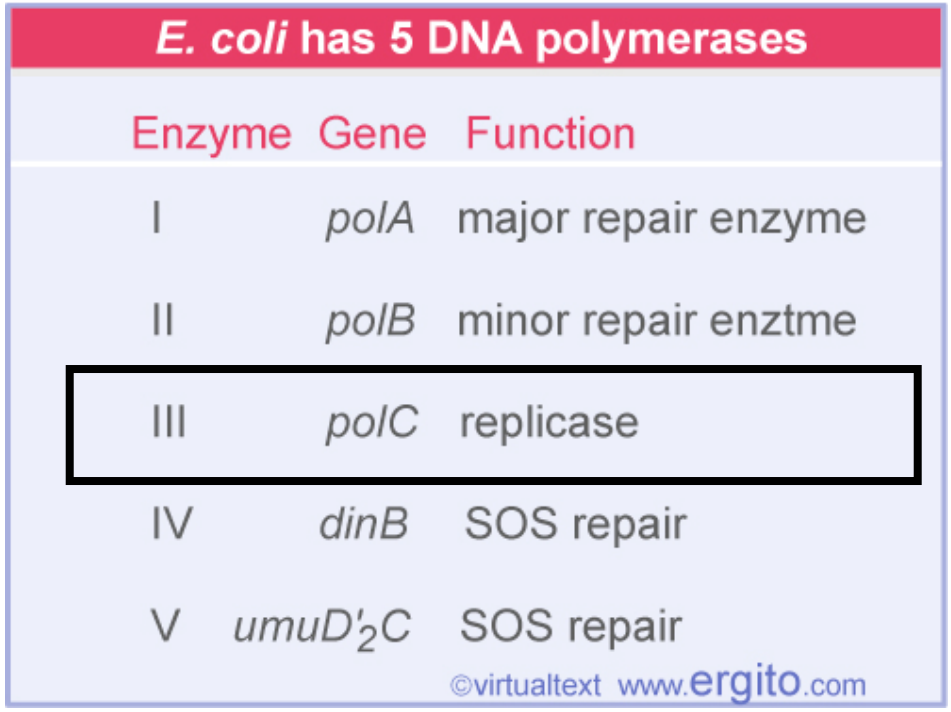
DNA is synthesized in both semi-conservative replication and in repair reactions

A bacterium or eukaryotic cell has several different DNA polymerase enzymes

One bacterial DNA polymerase undertakes semi-conservative replication; the others are involved in repair reactions

Eukaryotic nuclei, mitochondria, and chloroplasts each have a single unique DNA polymerase required for replication, and other DNA polymerases involved in ancillary or repair activities

# Organisms Contain Many Types of Polymerases




Multi-subunit:
*dnaE*-DNA synthetic activity
*dnaQ*- 3'-5' exo

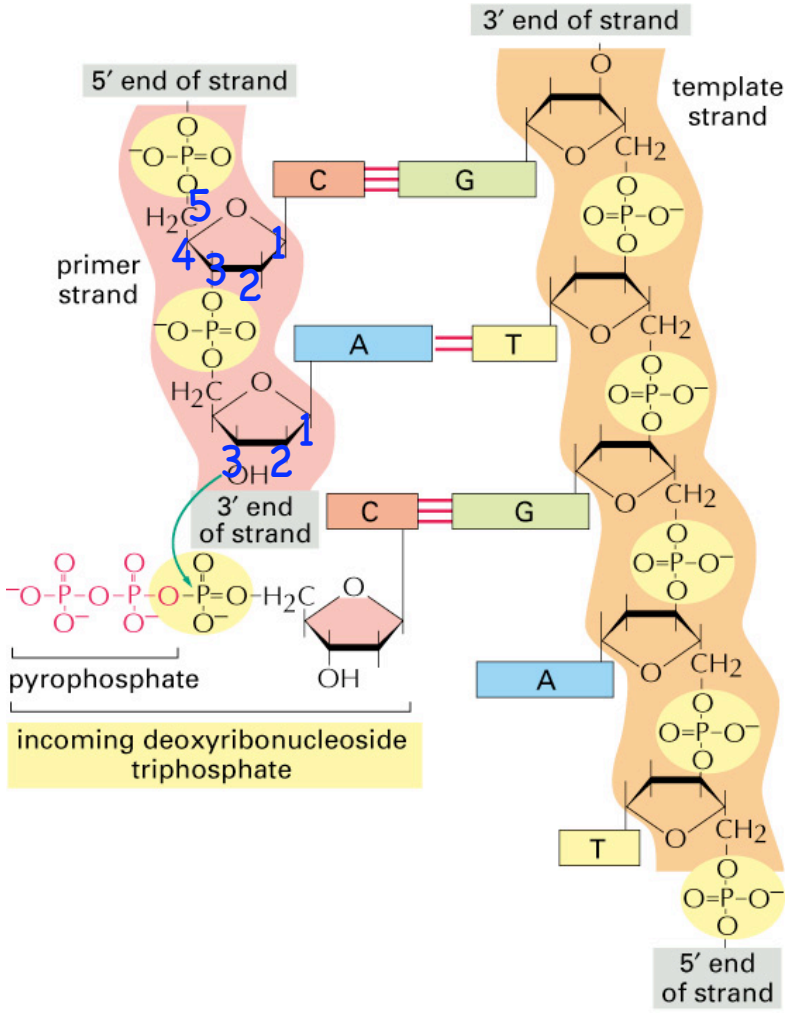
➢Pol III is the DNA replicase responsible for new strands of DNA in *E. coli*. Pol I also contributes, while others participate in specific DNA damage repair

# *E.coli* pol III part of a large assembly: Replicase

| Functional component | Subunit | Mass (kDa) | Gene | Activity |
|---|---|---|---|---|
| Core polymerase | α | 130 | *polC (dnaE)* | 5' to 3' polymerase |
| (aka, Pol III) | ε | 27.5 | *dnaQ (mutD)* | 3'-5' exonuclease |
| | θ | 10 | *holE* | Stimulates ε exonuclease |
| Gamma complex | τ | 71 | *dnaX* | Dimerizes cores |
| (Clamp loader/ | γ | 45.5 | *dnaX* | Binds ATP |
| ATPase) | δ | 35 | *holA* | Binds to β |
| | δ' | 33 | *holB* | Binds to γ and β |
| | χ | 15 | *holC* | Binds to SSB |
| | ψ | 12 | *holD* | Binds to χ and γ |
| Sliding clamp | β | 40.6 | *dnaN* | Processivity factor |

*Consider the polymerase…*

# Polymerases Perform the Fundamental Reaction in DNA Synthesis




The addition of a deoxyribonucleotide to the 3′ end of a primer strand

Figure 5–3. Molecular Biology of the Cell, 4th Edition.

Hoatlin Fall 2009

# DNA Synthesis

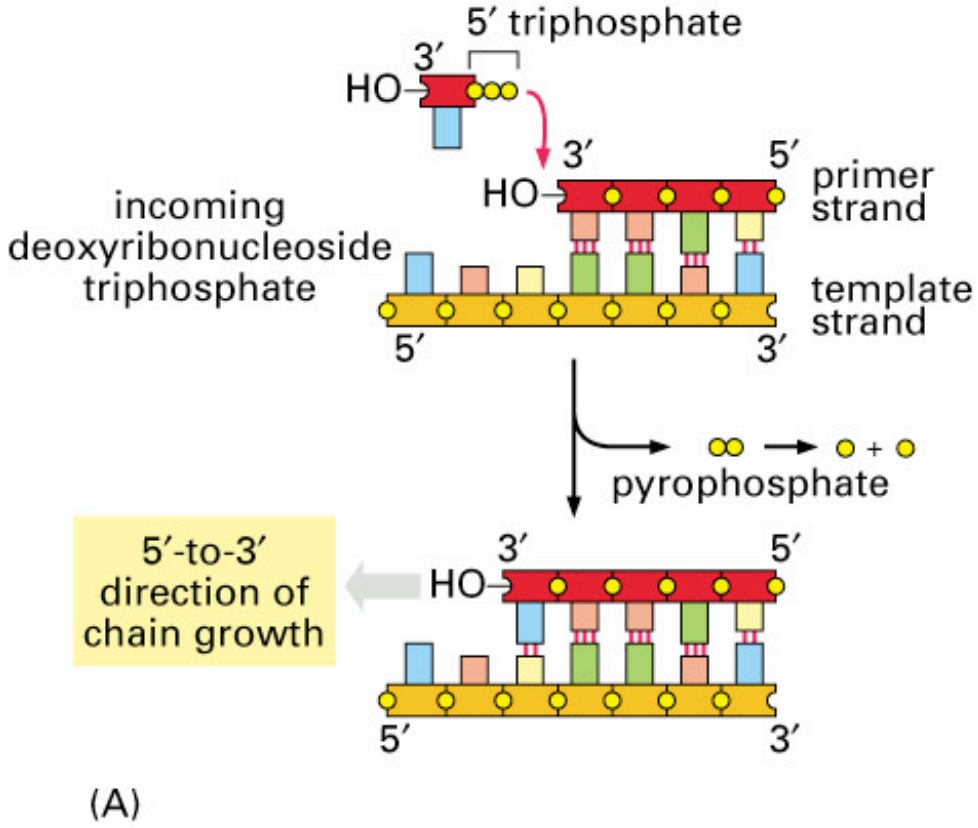



Figure 4–5 part 1 of 2. Molecular Biology of the Cell, 4th Edition.

Release of pyrophosphate followed by its hydrolysis to inorganic phosphate

# Features of a DNA Polymerase

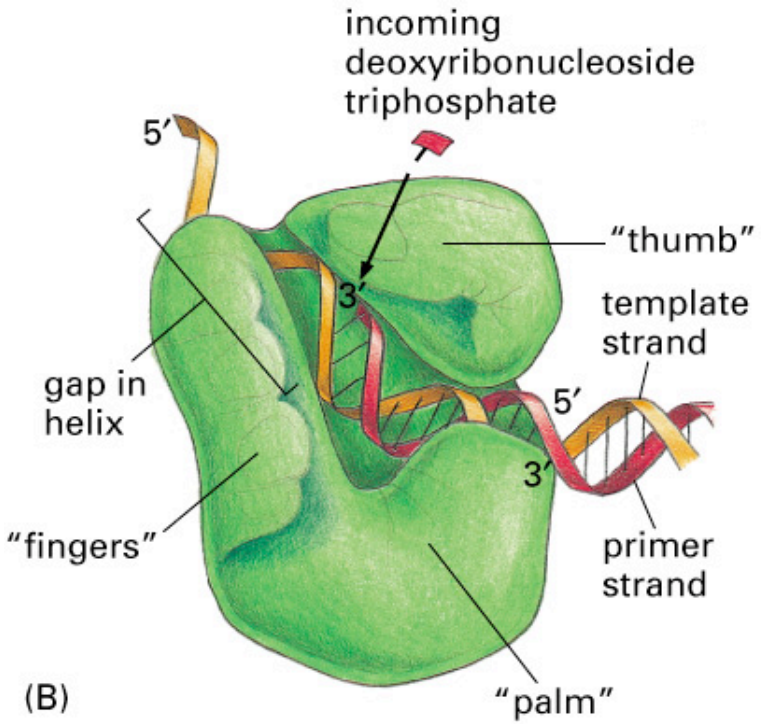



Figure 5–4 part 2 of 2. Molecular Biology of the Cell, 4th Edition.

Many DNA polymerases have a large cleft composed of three domains that resemble a hand
(if you use your imagination)

DNA lies across the "palm" in a groove created by the "fingers" and "thumb"-this is a good conceptual figure

DNA pols fall into 5 families based on sequence homologies. Palm is well conserved, thumb and fingers are analogous

# Replication Fork: Region of Active DNA Synthesis

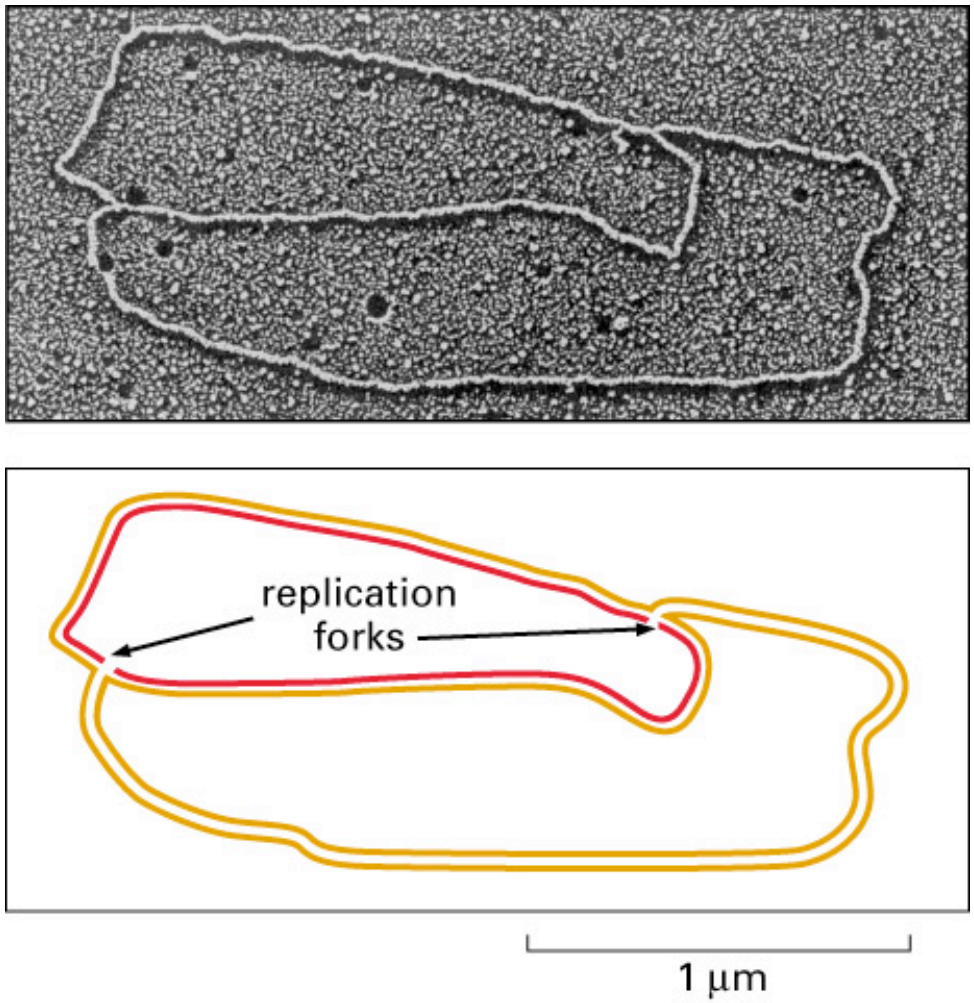
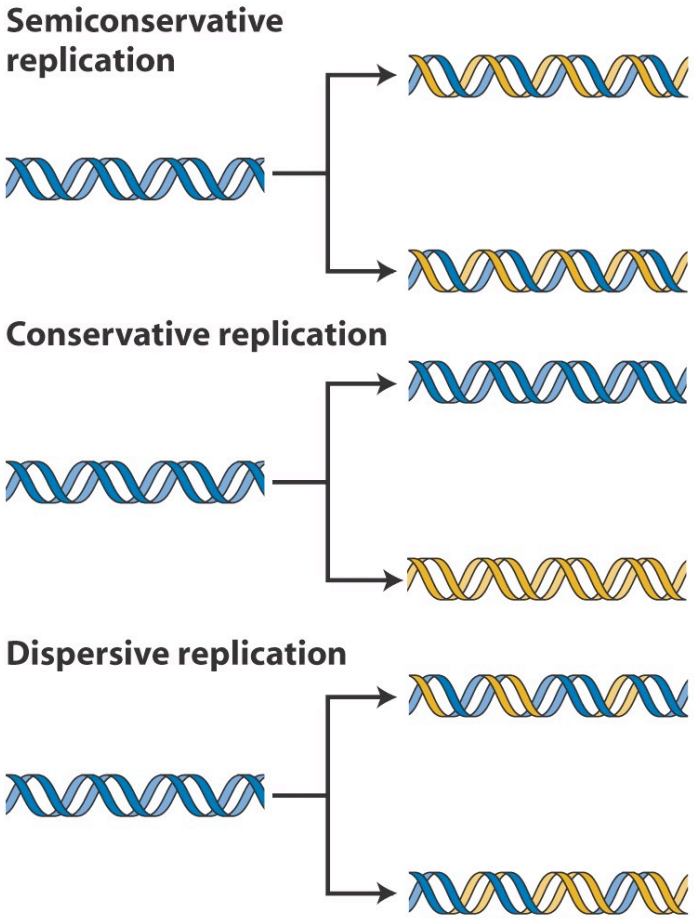



Figure 5–6. Molecular Biology of the Cell, 4th Edition.

DNA polymerases are in multi-enzyme complexes which synthesize the DNA of both daughter strands

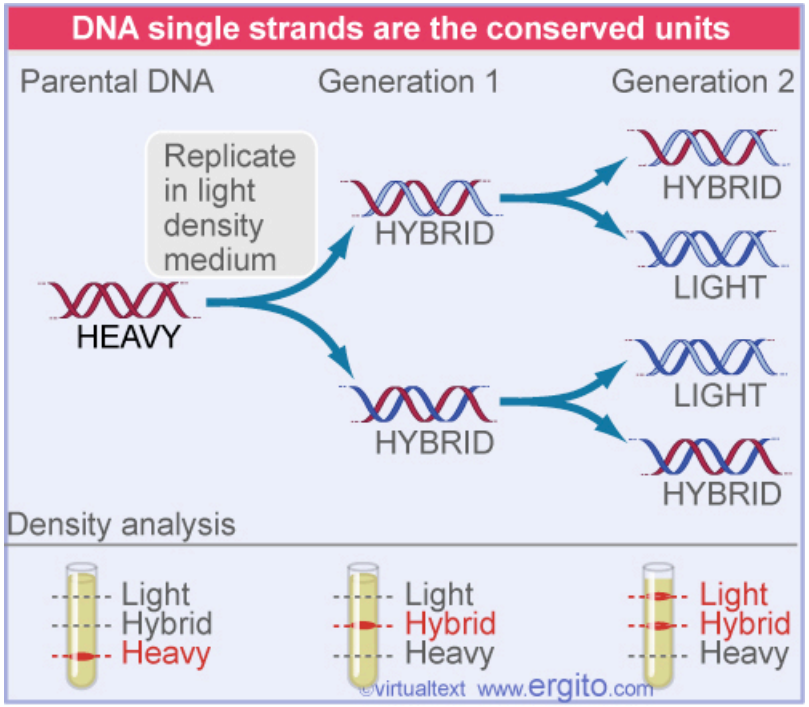
How are daughter stands synthesized?

# Possibilities for Replication Mechanism

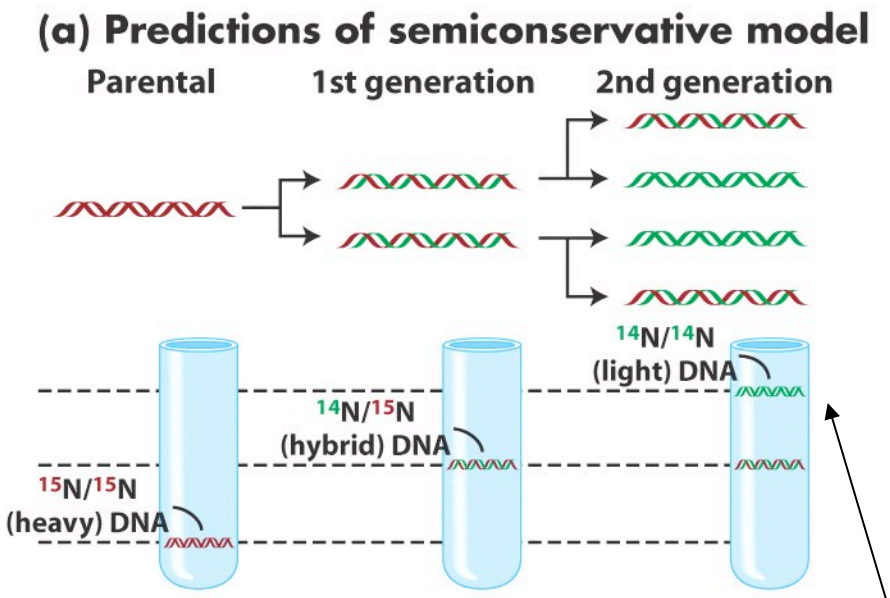


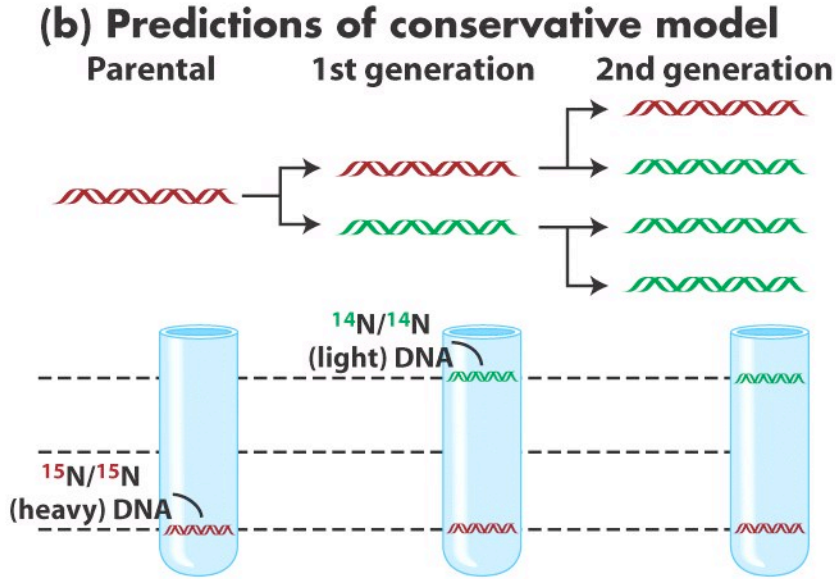

Hoatlin Fall 2009

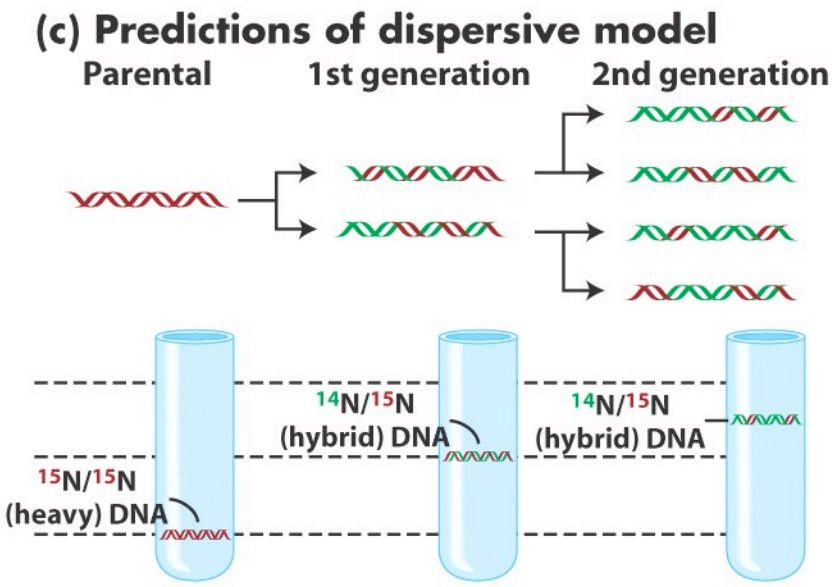
# Meselson-Stahl Experiment




- Parental DNA carries a "heavy" density label because organism grown in suitable isotope
- Switch to medium containing "light" isotope for replication
- Analyze replicated DNA in CsCl density gradients

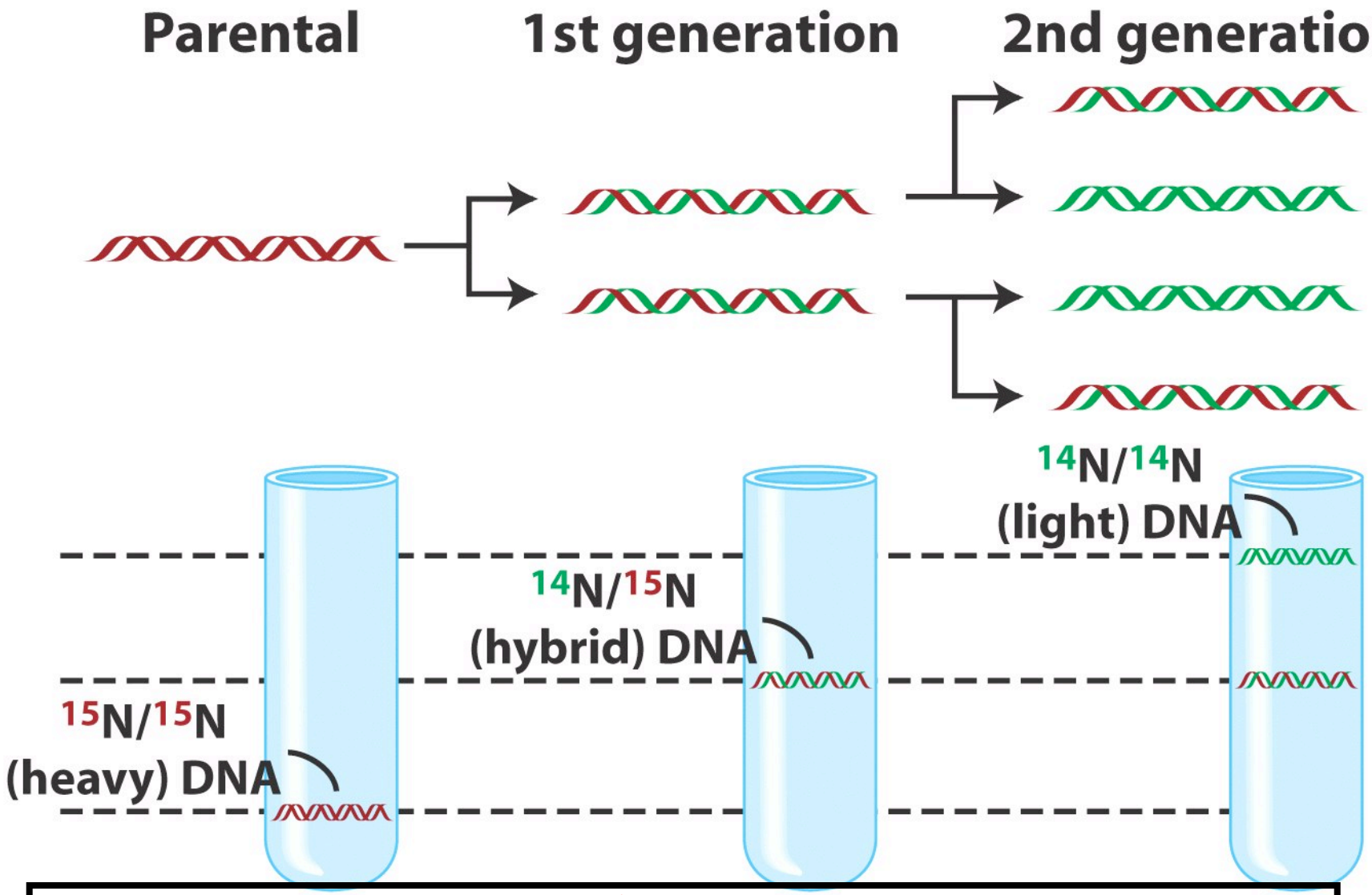



- Parental DNA consists of a duplex of two heavy strands
- Generation 1= one heavy parental strand and one light daughter strand
- Generation 2= half hybrid and half light

# Predictions of semiconservative model




This is what happens--but what is the mechanism?

# The Simple Model Fails

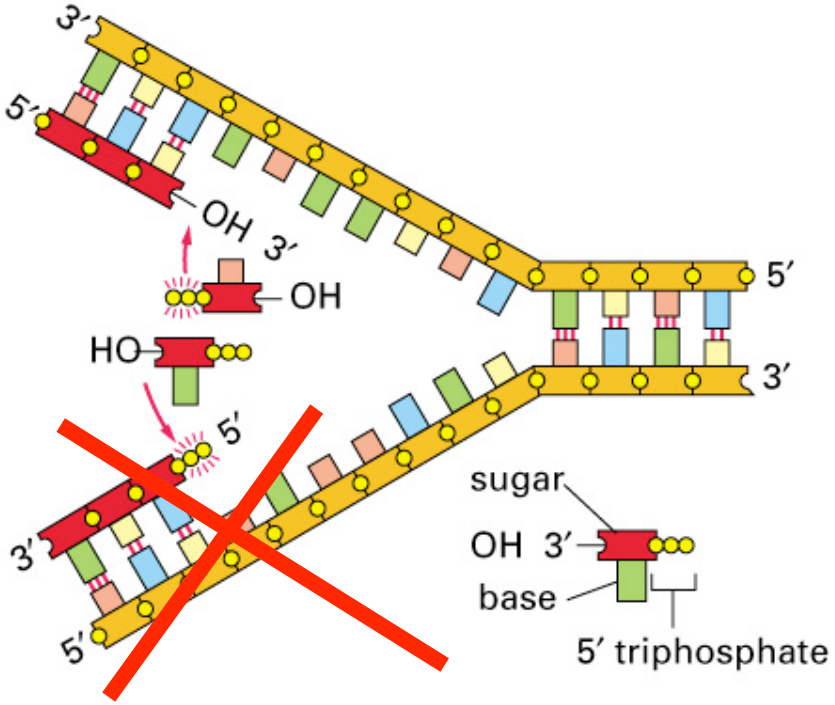



Figure 5–7. Molecular Biology of the Cell, 4th Edition.

Exactly how are nucleotides added to the anti-parallel strands?

Requires one strand to grow in 5′-to-3′ direction, while the other strand grows in the 3′-to-5′ direction

But DNA polymerase requires an incoming nucleotide in a 5′-to-3′ direction

# The Replication Fork is Asymmetric

Leading strand: synthesized *almost* continuously

Lagging strand: discontinuous synthesis. Okazaki fragments are polymerized in 5′-to-3′ direction and stitched together after synthesis. Lagging, or delayed, because it must wait for leading strand to expose template to synthesize Okazaki fragment (more on this later)

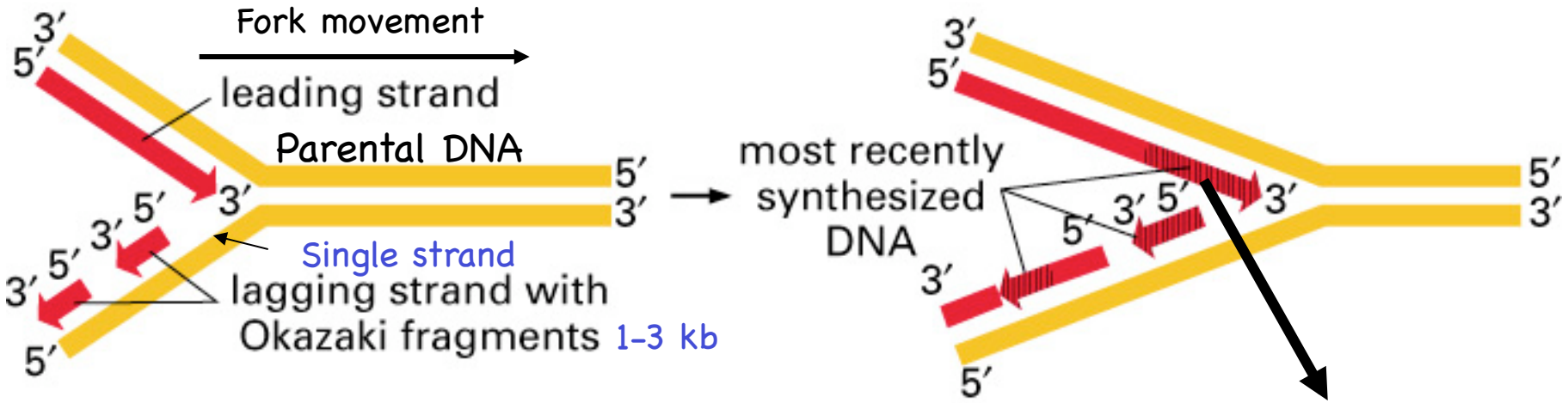



Figure 5–8. Molecular Biology of the Cell, 4th Edition.

# Hi-Fi Replication Facts

- During synthesis mispairing is inevitable
- Correct nuc has a higher affinity for moving polymerase compared to incorrect--e.g., correct base-pairing, correct shape
- Incorrectly bound nucleotide more likely to dissociate as polymerase undergoes conformational change after binding but before covalent addition of nucleotide to growing chain

# More Hi-Fi: Proofreading by DNA Polymerase During Replication

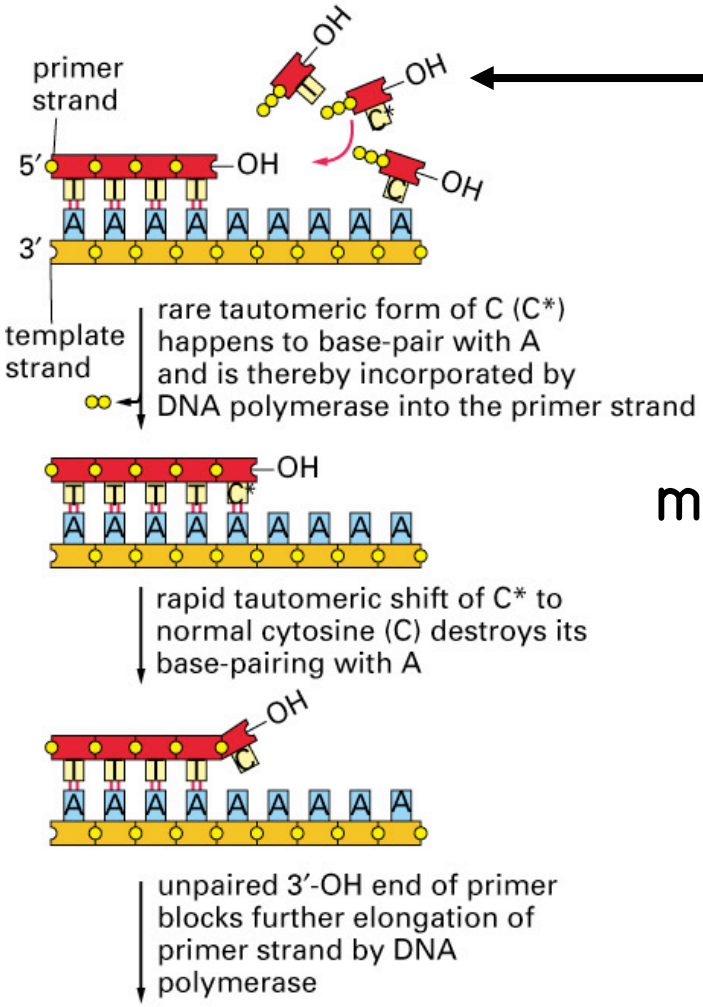



Figure 5–9 part 1 of 2. Molecular Biology of the Cell, 4th Edition.

mis-incorporation

Note: Polymerase requires a *base-paired* 3′-OH terminus

# Cont'd---The Hi-Fi DNA Polymerase

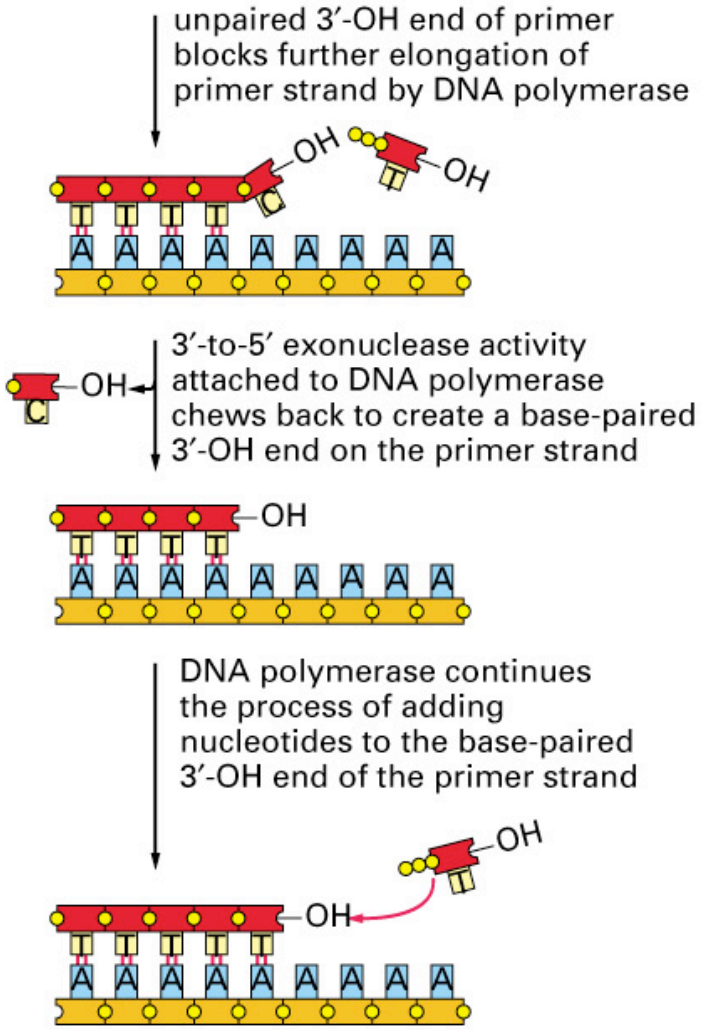



Figure 5–9 part 2 of 2. Molecular Biology of the Cell, 4th Edition.

A separate catalytic site (3′-5′ exonuclease) in the polymerase-- or in a sub-unit-- clips off unpaired residues until a base-paired terminus is regenerated

Fidelity of replication is improved by proofreading by a factor of approx. 100

All bacterial Pols have 3′-5′ exo activity

# DNA Polymerase: Two Modes

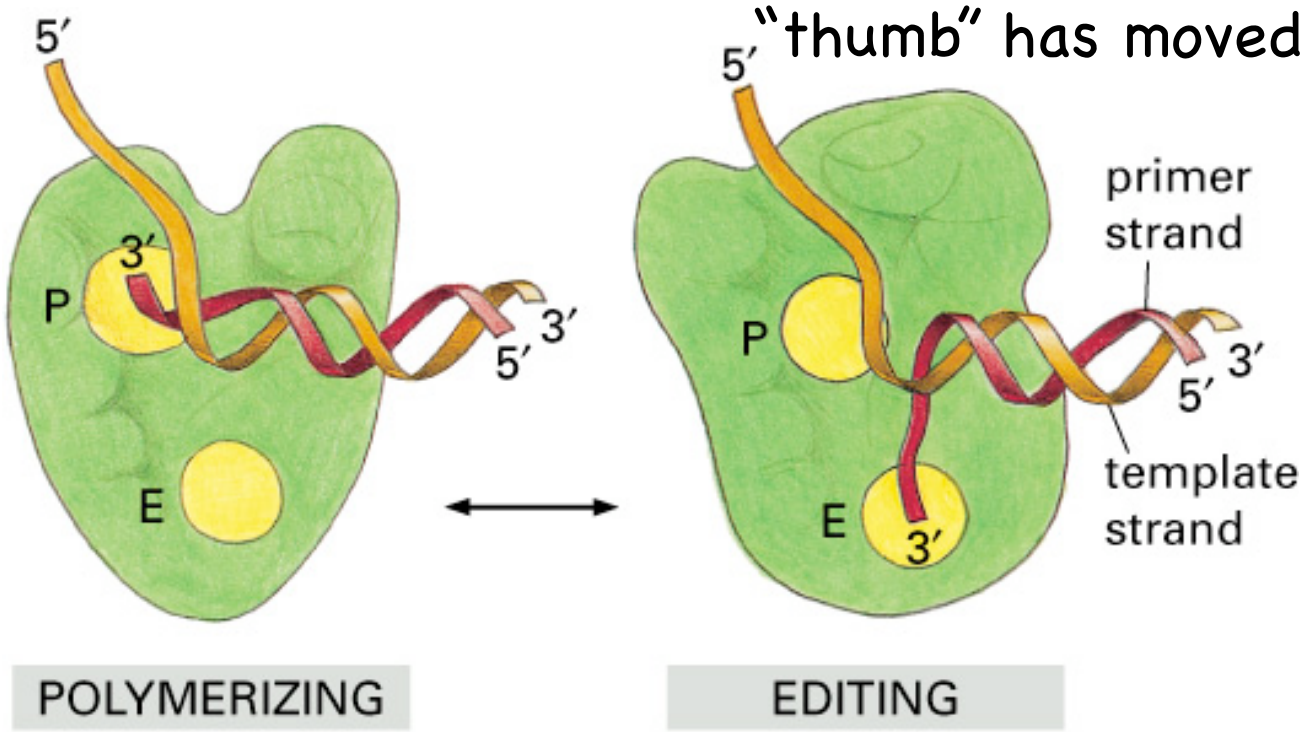
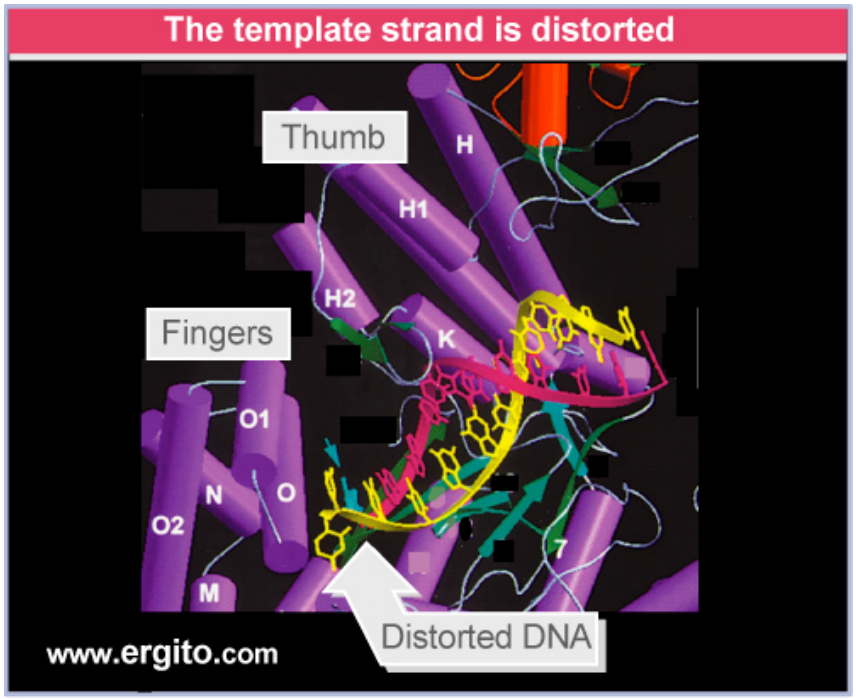



Figure 5–10. Molecular Biology of the Cell, 4th Edition.

P: polymerization site
E: exonucleolytic site

# Example: Polymerase from T7 Bacteriophage



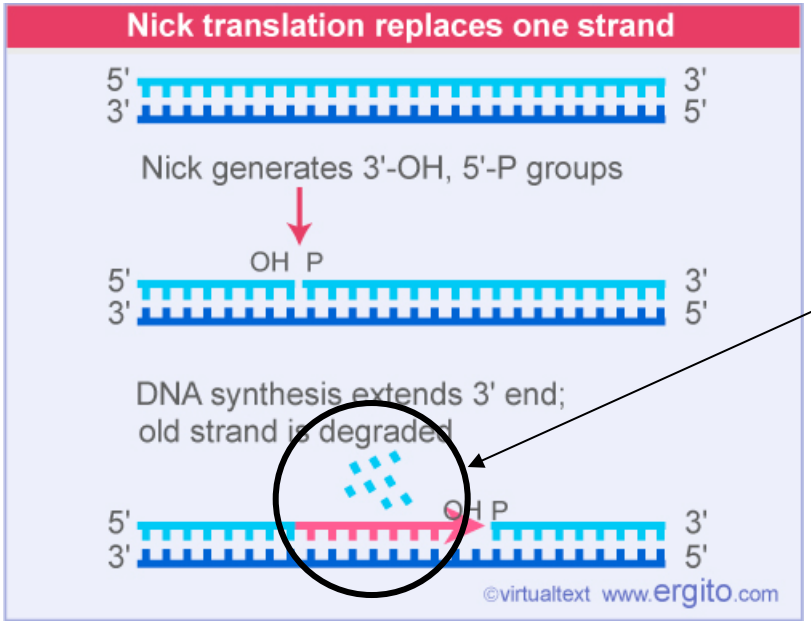

Template strand takes a sharp turn so it can be exposed to the incoming nucleotide

Inward rotation of the O helix grasps incoming nuc and creates a catalytic site by rotation of fingers and thumb.

Cyclic: reversed when nuc is incorporated into growing DNA strand--and then nuc moves out of site

When mismatched bp is in catalytic site, fingers cannot rotate toward the palm, leaving 3' end free to bind active site in exonuclease domain

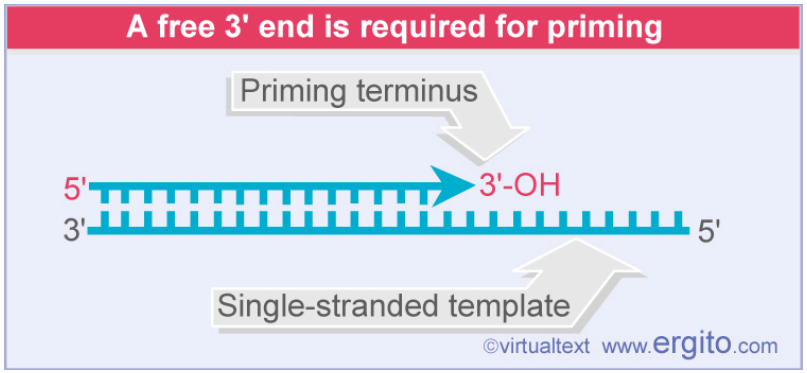
# DNA Pol I Nick Translation*




- DNA pol I, 103 kD
- Cleavage into two domains:
  - 68 kD (Klenow fragment) with pol and 3′-5′ exonuclease activities
  - 35 kD fragment with 5′-3′ exonucleolytic activity.
- These 3 functions give DNA pol I unique ability to start replication *in vitro* at a nick in DNA
- The nick just moves along (is translated..kind of an unfortunate description in this case)

*very useful at the bench!

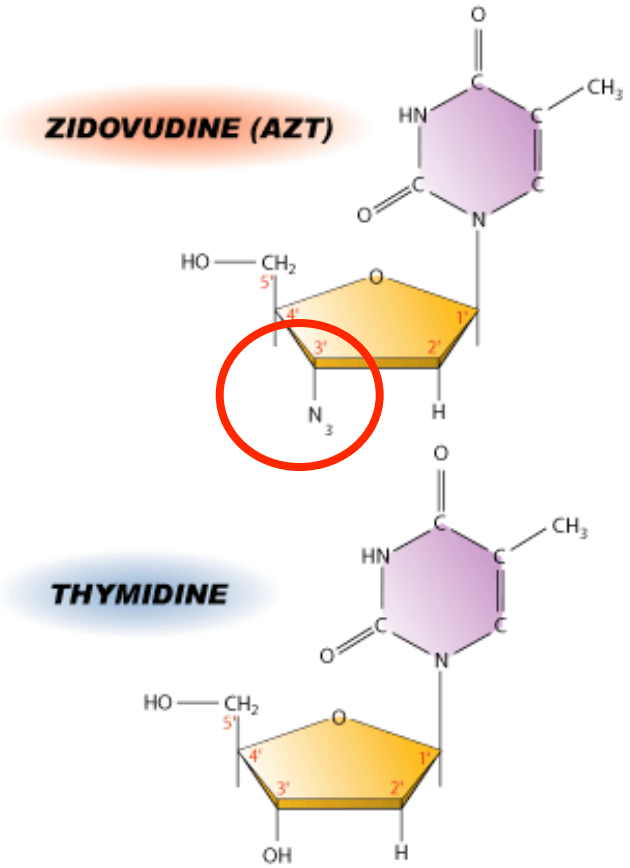
# DNA Pol's Needs are Rigid

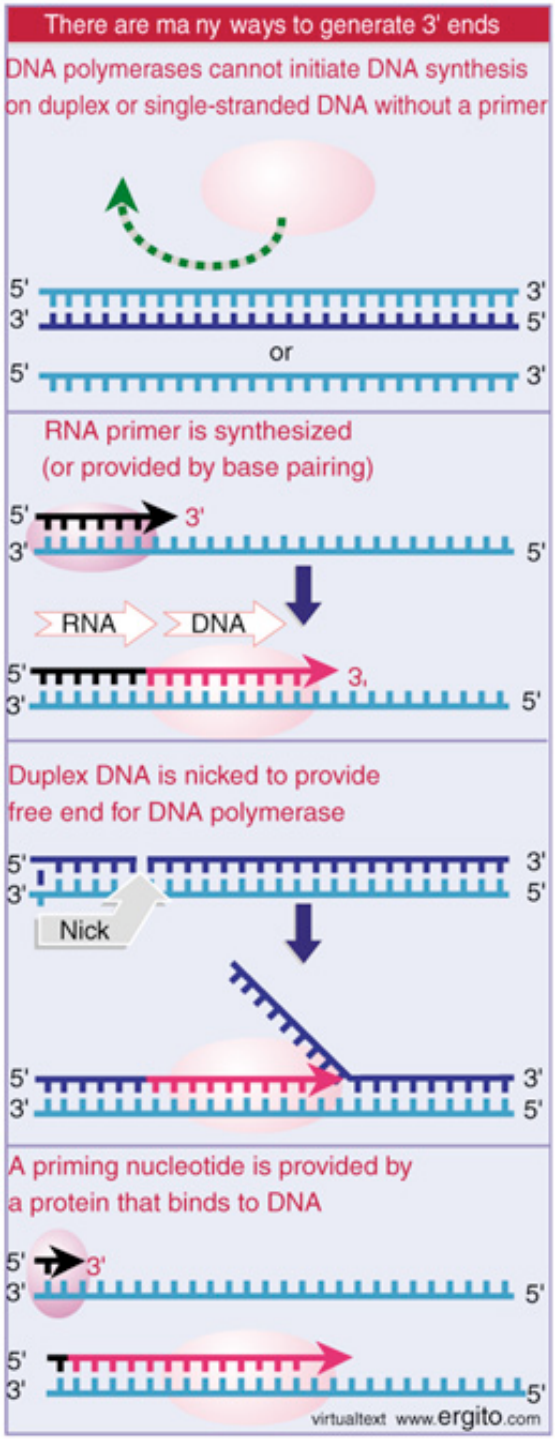



- All DNA polymerases require a 3´-OH priming end to initiate DNA synthesis
- The priming end can be provided by an RNA primer, a nick in DNA, or a priming protein
- Template strand must be single stranded

# Rational Strategy for an anti-HIV Drug

- Azidodeoxythymidine (AZT) is an analog of thymidine
- AZT is a *reverse transcriptase* inhibitor
- Reverse transcriptase enzymes are polymerases required for retroviral replication
- The N3 group does not allow DNA synthesis to continue (why?)
- 100x to 300x affinity for rt vs DNA pol. Gamma DNA pol (mitochondrial pol) is affected with associated side effects






# 3 Ways To Satisfying Pol's Need for a 3' End

RNA primer: commonly used in cellular DNA replication and some viruses (pre-formed RNA to prime reverse transcription of RNA)--details on next slide

Nicking: primer terminus generated within duplex DNA. Pre-existing strand displaced by new strand (not degraded, as in "nick translation" -- only pol I can do that)

Protein presents a nucleotide to DNA pol (adenoviral strategy)

Hoatlin Fall 2009

# DNA Primase: RNA Primer Synthesis

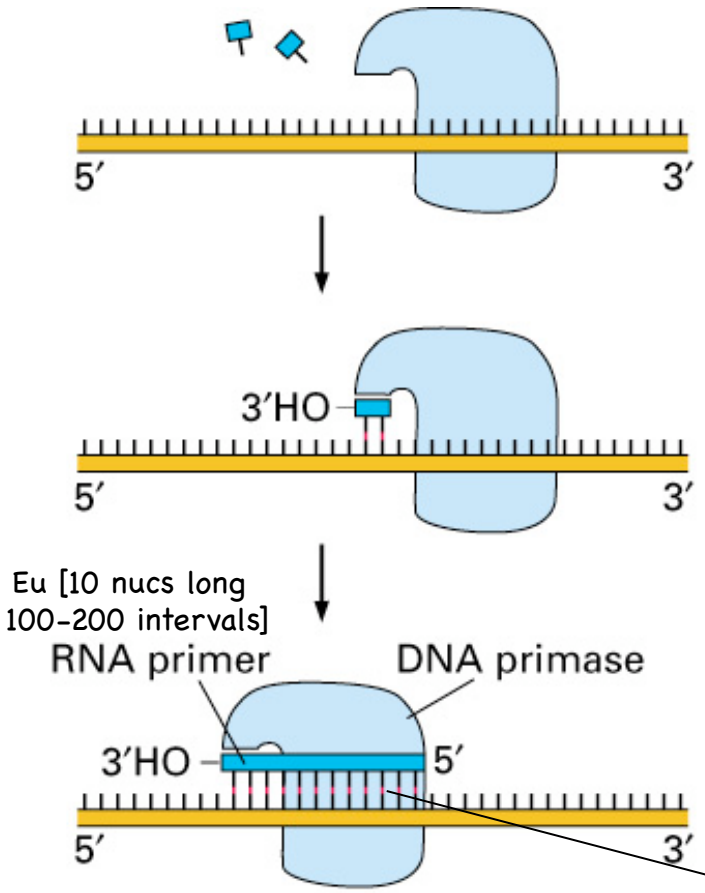



Figure 5–12. Molecular Biology of the Cell, 4th Edition.

Synthesizes a short polynucleotide (approx 10) in the 5′-to-3′ direction and then stops

Uses DNA as a template and builds primer from ribonucleoside triphosphates

Unlike DNA pol, which requires a primer, DNA primase can start up a new template-guided polynucleotide strand

Product of the *dnaB* gene (remember the dna mutants? Would dnaB be quick or slow stop at non-permissive temp?)

RNA-DNA hybrid

# Lagging Strand Synthesis- Many RNA Primers

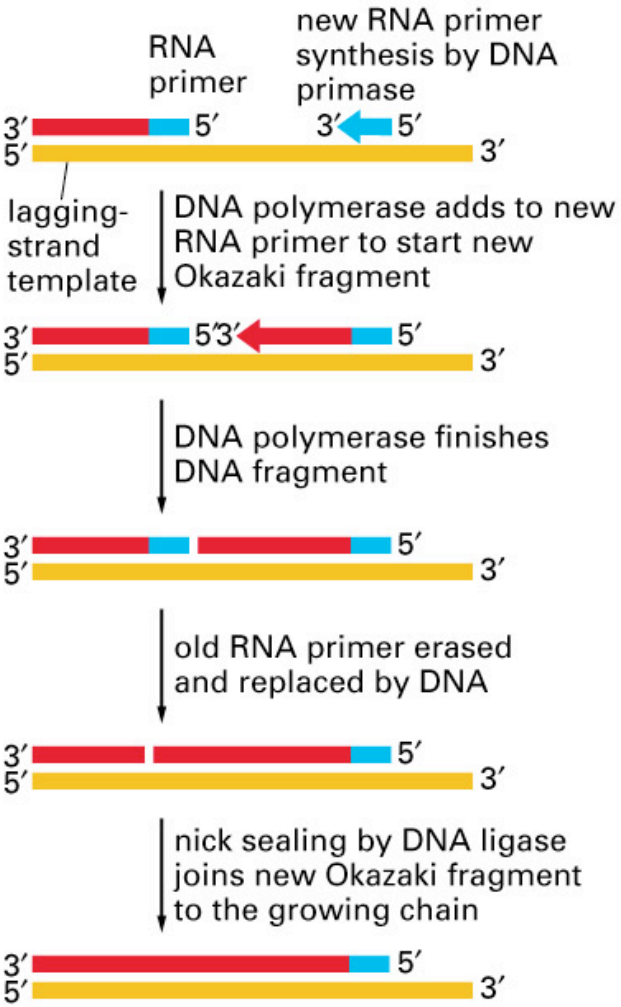



Figure 5–13. Molecular Biology of the Cell, 4th Edition.

RNA primers have a free 3′-OH, substrate for DNA pol

RNA removed, DNA filled in

Nicks between Okazaki fragments sealed by DNA ligase

# Helicases: Opening up DNA Double Helix in Front of the Replication Fork

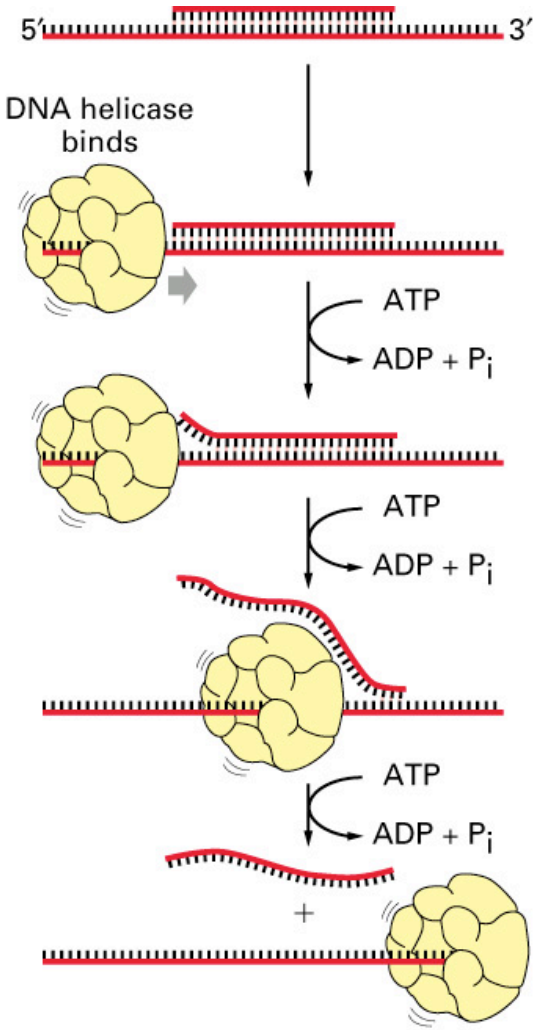
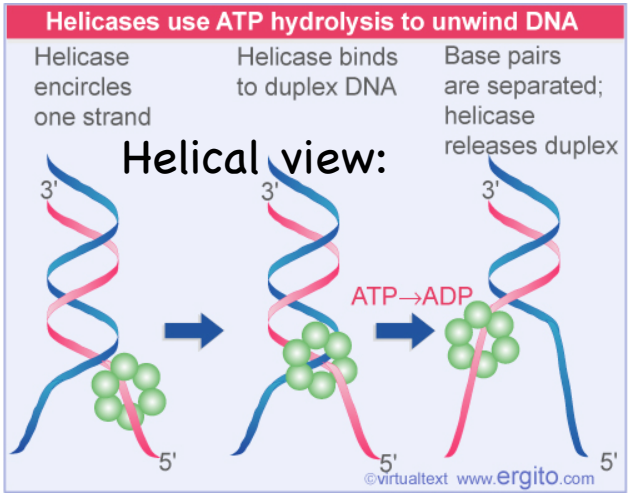



Figure 5–15. Molecular Biology of the Cell, 4th Edition.

Hydrolysis of ATP changes shape of helicase in cyclical manner that allows helicase to propel rapidly (1000 nucleotide pairs/sec) along a DNA single strand, prying apart the helix

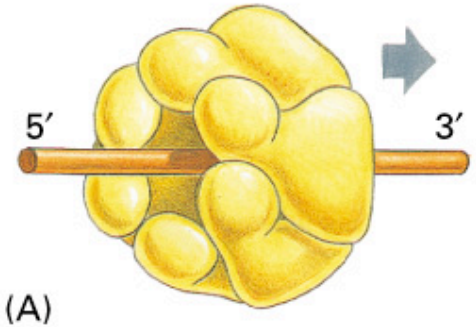
12 different helicases in *E. coli*

DnaB encircles the lagging strand

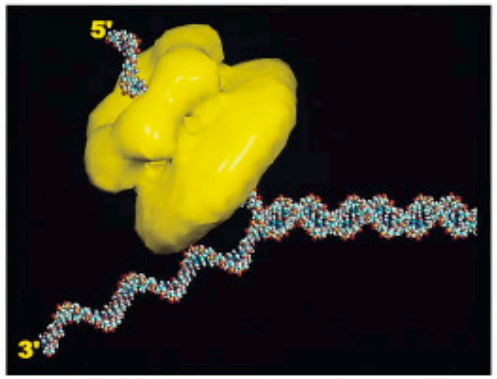


# Structure of a DNA Helicase

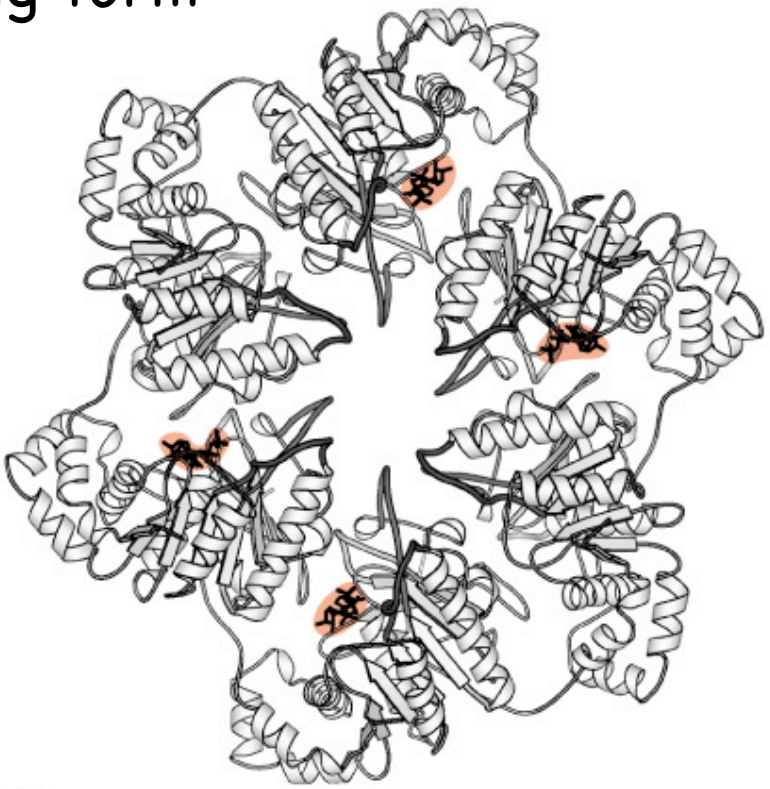
Hexamer in a ring form




(A)




(B)



(C)

Bacteriophage T7 replicative helicase (red is ATP). Six identical s.u. bind single strand and duplex DNA alternately and hydrolyze ATP in an ordered fashion to propel molecule along DNA-- a single strand that passes through the central hole

Replication fork and helicase to scale

nature Vol 442|20 July 2006|doi:10.1038/nature04943

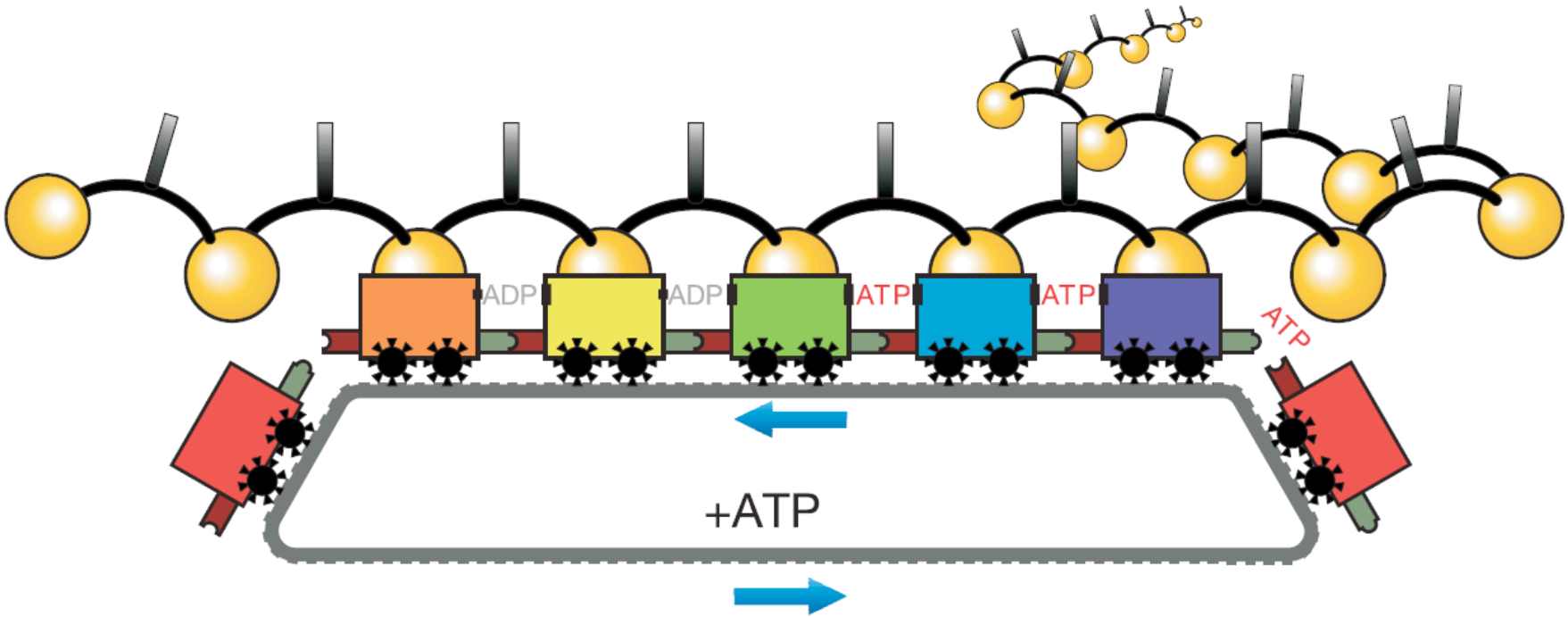
ARTICLES

# Mechanism of DNA translocation in a replicative hexameric helicase

Eric J. Enemark[1] & Leemor Joshua-Tor[1]

- Crystal structure of hexameric helicase (viral)
- Showed the path of DNA; Each subunit of the hexamer contributes a nucleotide binding loop that binds single-stranded DNA on one side of the ring and escorts it through the channel like an escalator
- Five adjacent subunits escort five adjacent nucleotides and the loop of the sixth subunit is recycled from the exit side to the entry side to pick-up a new nucleotide.
- the hexamer encircles single-, and not double-, stranded DNA.

# Escort Model




Hoatlin Fall 2009

# Heicase Helpers: Single strand DNA-binding (SSB) Proteins

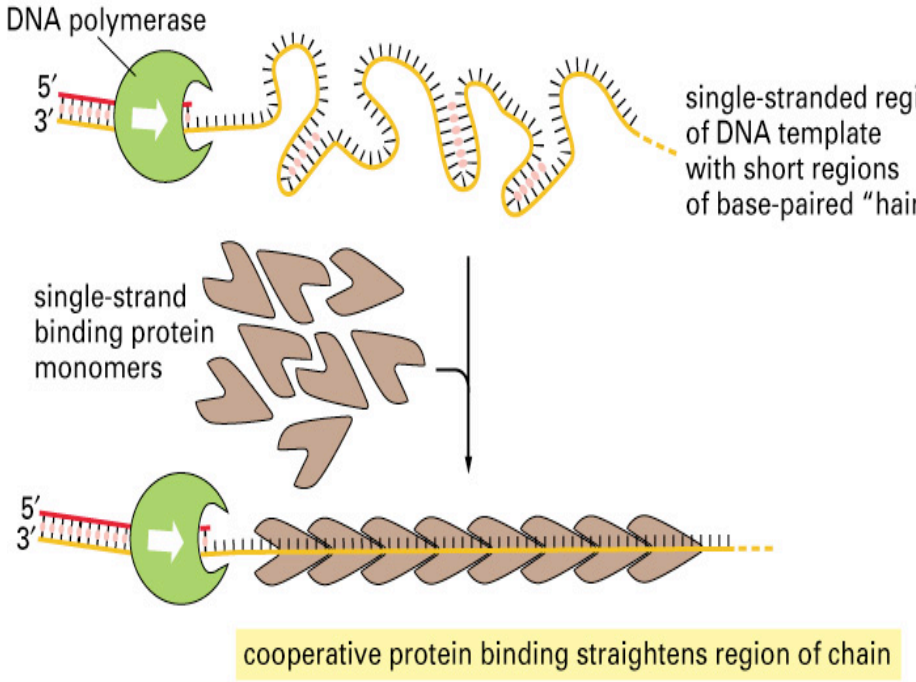



Figure 5–17. Molecular Biology of the Cell, 4th Edition.

Bind cooperatively to exposed single strands of DNA

SSBs do not interfere with base-pairing

Unable to open helix directly

Aid helicases by stabilizing unwound DNA-reduce "hairpins"

*ssb* mutants have a quick-stop phenotype-defective in repair and recombination as well as in replication

# Structure of a Single-Strand Binding Protein (human RPA)

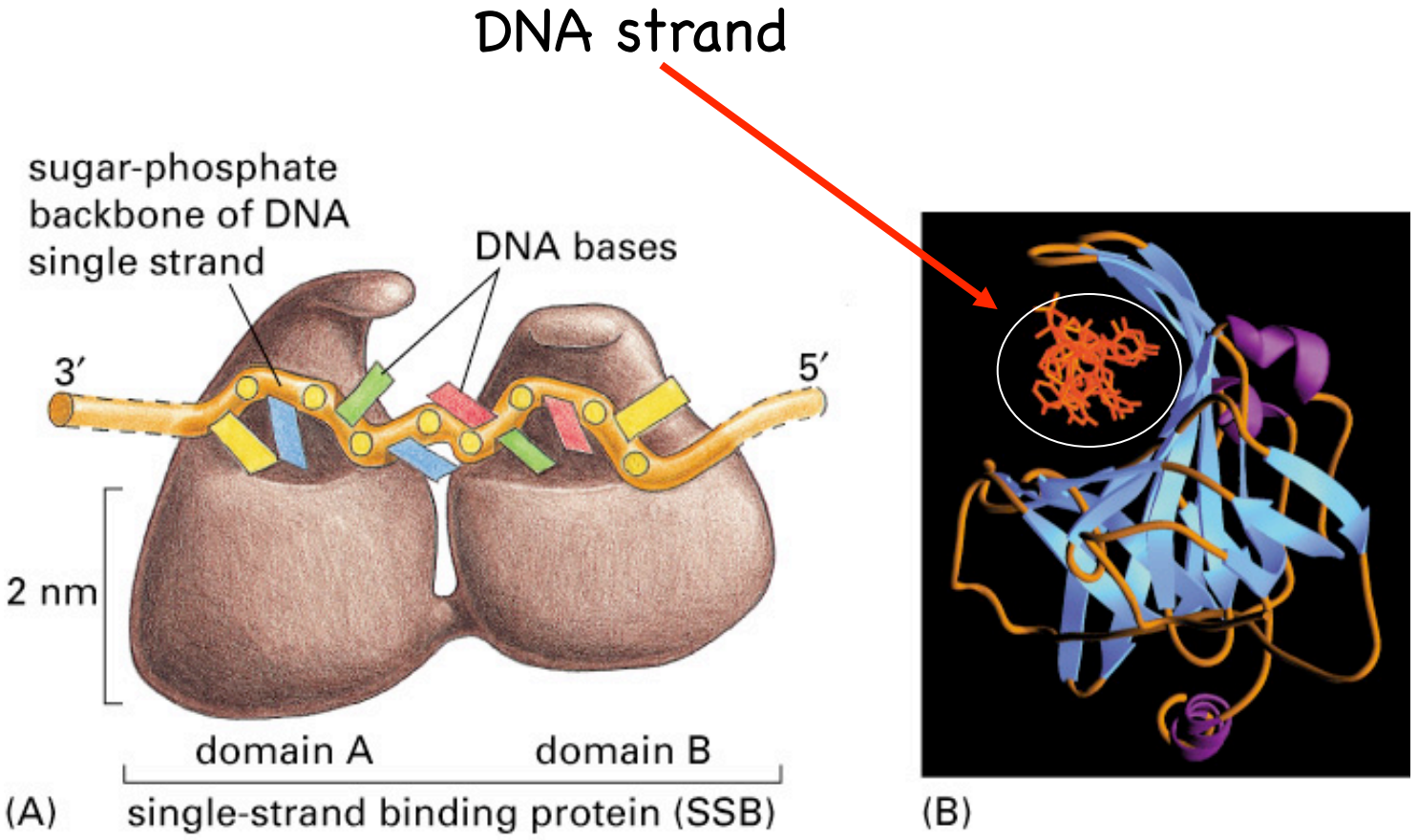
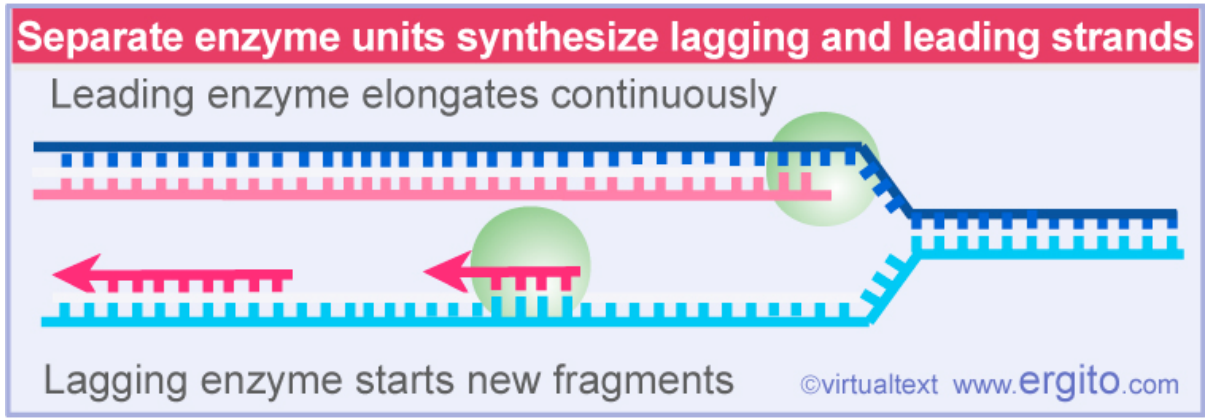



Figure 5–18. Molecular Biology of the Cell, 4th Edition.

# Coordinating Synthesis of Leading and Lagging Strands




- Different enzyme units are required to synthesize the leading and lagging strands
- In *E. coli* both these units contain the same catalytic subunit (DnaE), supported by other proteins that differ between leading and lagging strands
- In other organisms, different catalytic subunits may be required for each strand

# Replication at the Fork


- "hybrid" of prokaryote and eukaryote replication proteins and dynamics

http://www.wehi.edu.au/education/wehi-tv/dna/replication.html

# *E.coli* pol III part of a large assembly: Replicase

| Functional component | Subunit | Mass (kDa) | Gene | Activity |
|---|---|---|---|---|
| Core polymerase | α | 130 | *polC (dnaE)* | 5' to 3' polymerase |
| (aka, Pol III) | ε | 27.5 | *dnaQ (mutD)* | 3'-5' exonuclease |
| | θ | 10 | *holE* | Stimulates ε exonuclease |
| Gamma complex | τ | 71 | *dnaX* | Dimerizes cores |
| (Clamp loader/ | γ | 45.5 | *dnaX* | Binds ATP |
| ATPase) | δ | 35 | *holA* | Binds to β |
| | δ' | 33 | *holB* | Binds to γ and β |
| | χ | 15 | *holC* | Binds to SSB |
| | ψ | 12 | *holD* | Binds to χ and γ |
| Sliding clamp | β | 40.6 | *dnaN* | Processivity factor |

*Consider the polymerase…*

# DNA polymerase III is a holoenzyme with 3 subcomplexes

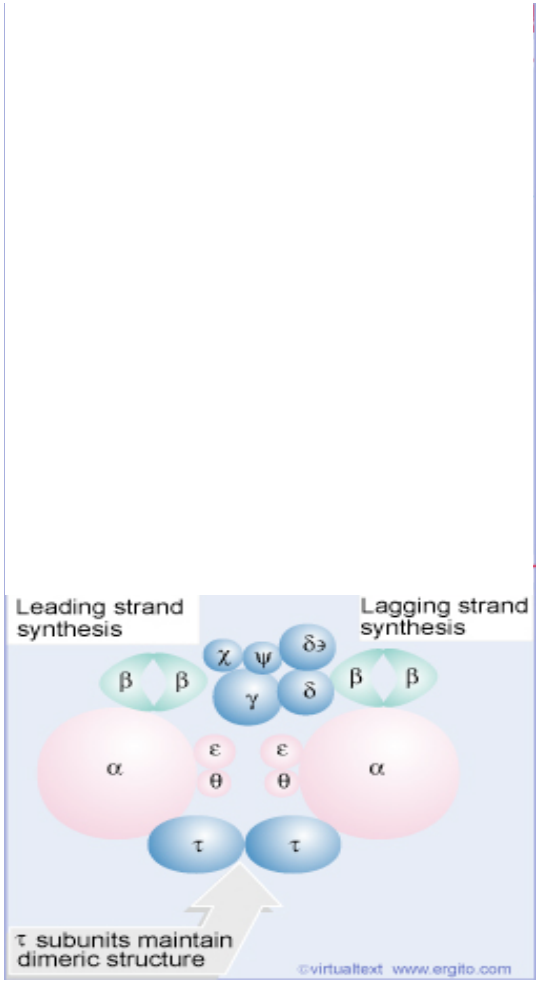
The *E. coli* replicase DNA polymerase III is a 900 kD complex with a dimeric structure

Each monomeric unit has a catalytic core, a dimerization subunit, and a processivity (stays on DNA) component

A clamp loader places the processivity subunits on DNA, and they form a circular clamp around the nucleic acid

One catalytic core is associated with each template strand

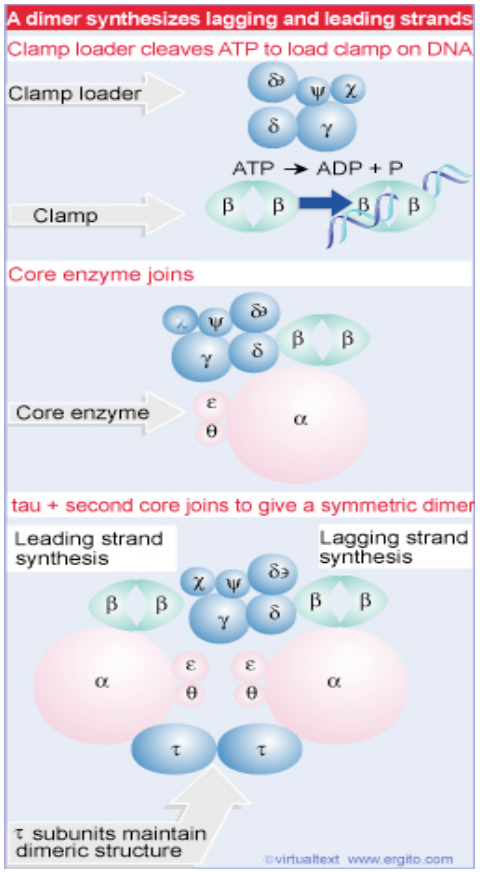
# E. Coli Replicase (Pol III)




10 proteins:

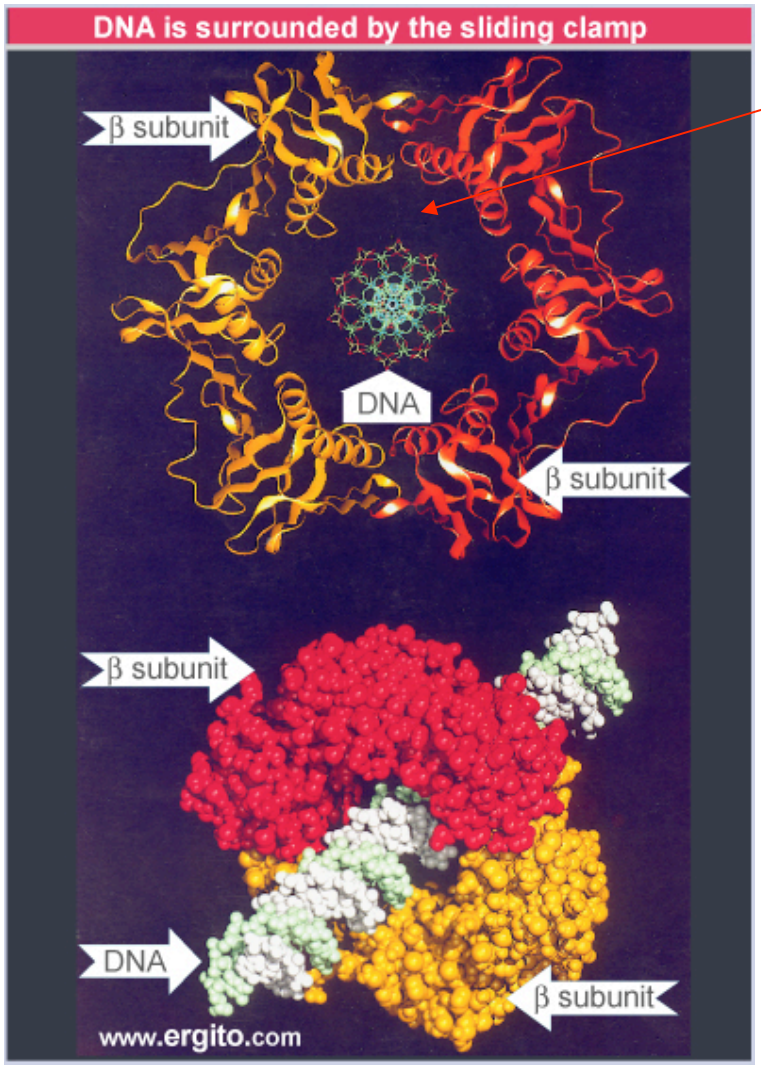
- 2 copies of the catalytic core, each containing the $\alpha$ subunit (pol activity), $\varepsilon$ su (3'-5' proofreading exonuclease), and $\theta$ su (stimulates exonuclease)
- Two copies of dimerizing s.u., $\tau$, which link the catalytic cores
- Two copies of the processivity su, $\beta$, clamps which are responsible for holding the catalytic cores onto the template strands
- The $\gamma$ complex (the clamp loader), five proteins that place the processivity su on DNA

# DNA pol III Assembles in Stages




- The β dimer (clamp) plus a γ complex (clamp loader) recognizes the primer-template and forms pre-initiation complex. Transfer of β su to primed template, Clamp formed by β su
- Binding to DNA changes conformation of site on β that binds to γ. Now β has high affinity for core pol, and core is brought in to DNA
- The τ dimer binds to core pol and provides a dimerization that allows second core (and β clamp) to bind

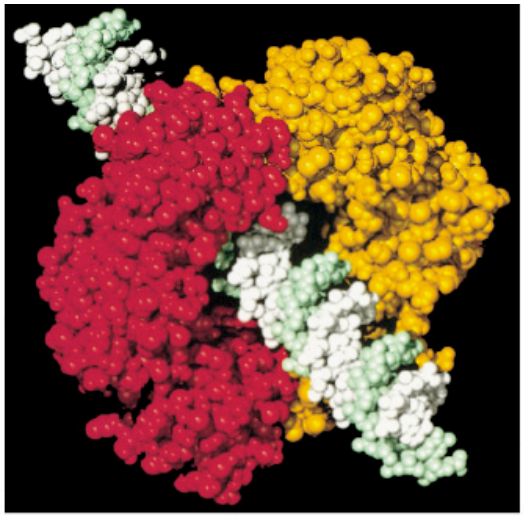
# The Sliding Clamp



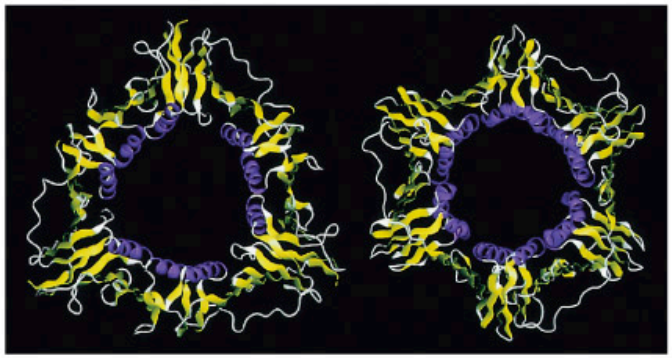

Water slide

The core polymerase on the leading strand is processive because the sliding clamp keeps it on the DNA
-ring shaped

# Structure of Clamp Proteins



(A)



(B)

Figure 5–19 part 1 of 2. Molecular Biology of the Cell, 4th Edition.

Clamp protein from *E. coli*, helix shown through central hole (E. coli has four different clamp loader proteins)

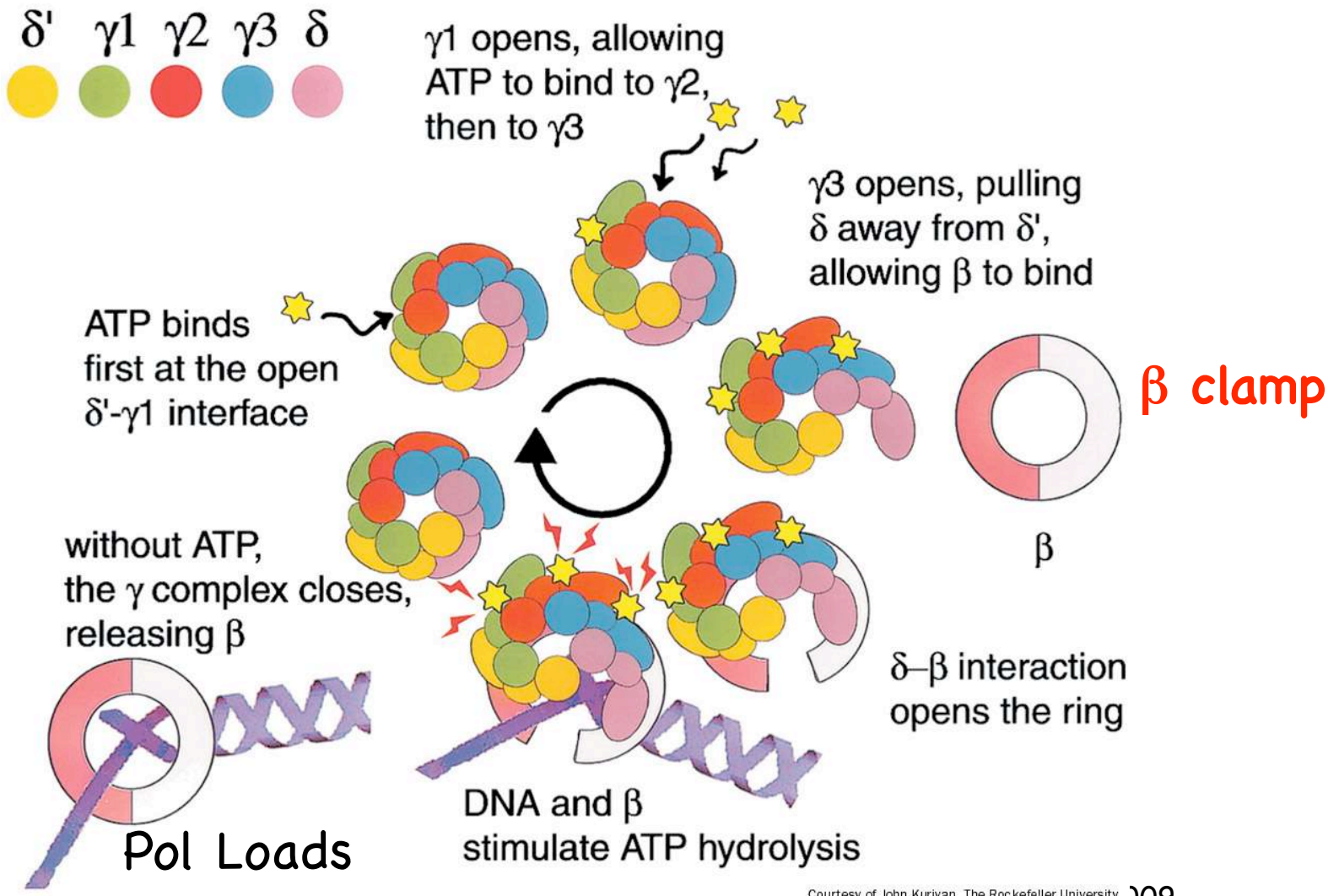
Clamp Protein (β clamp) from *E. coli* (left) and human clamp protein, PCNA (right)--very similar

-

Clamps interact with recombination, repair, and cell cycle proteins

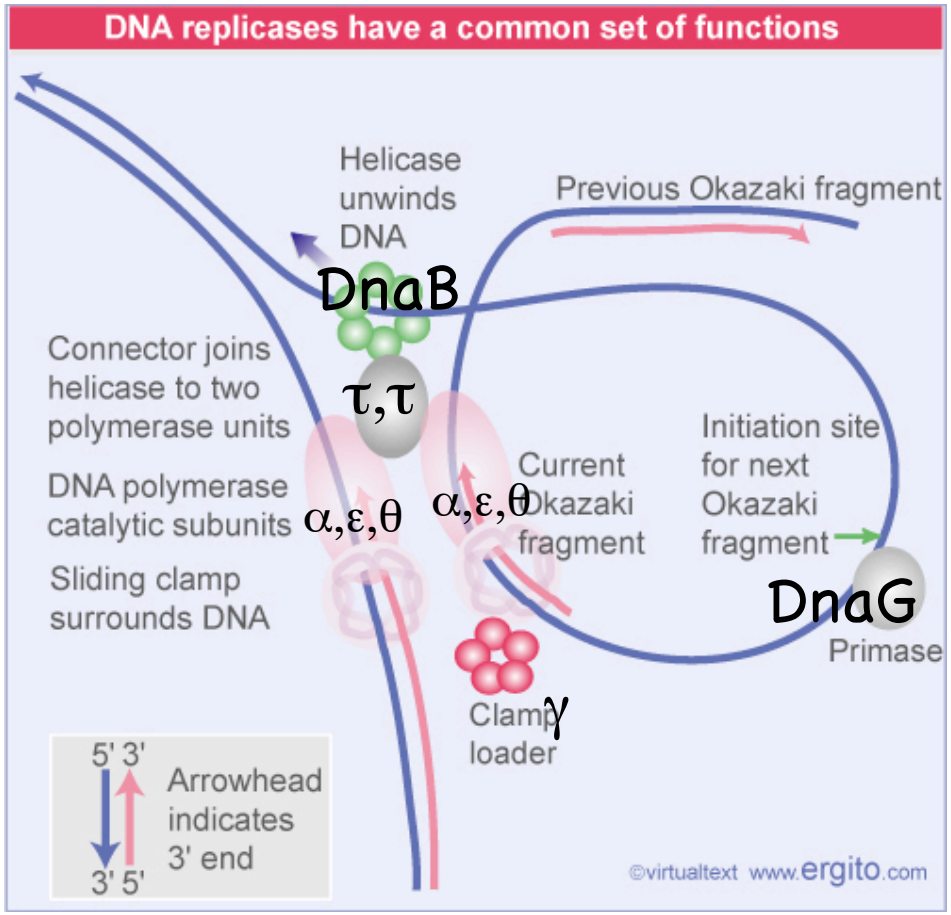
-

# Clamp Loader Cycle (*E. coli*)




Courtesy of John Kuriyan, The Rockefeller University

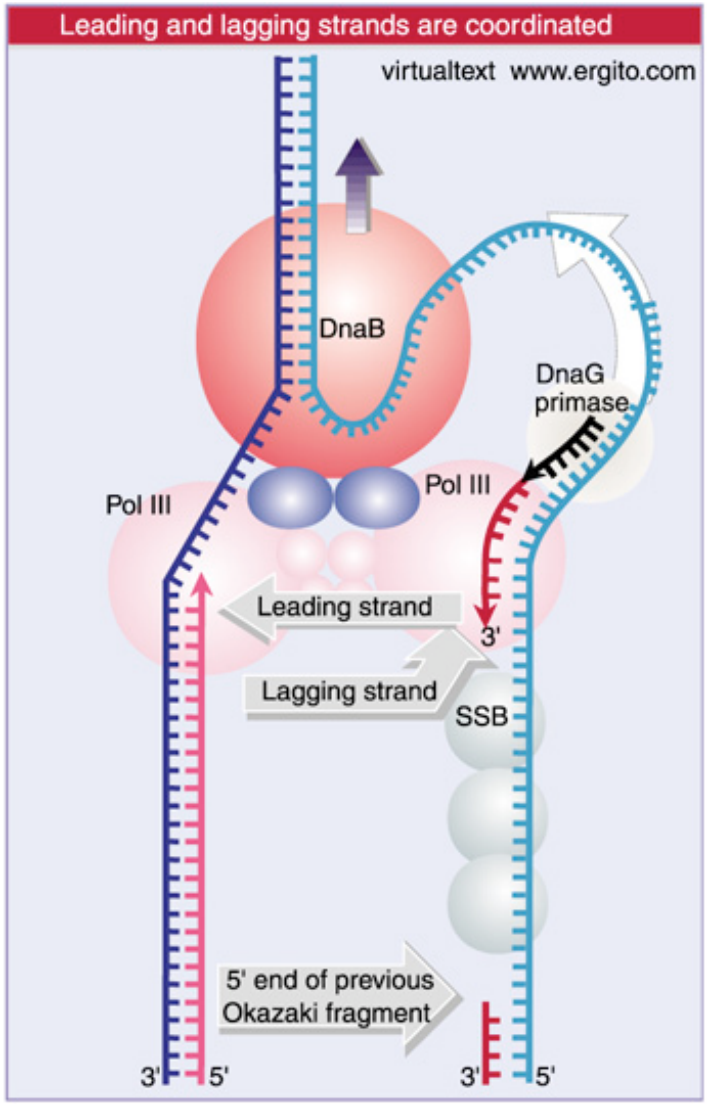
# A Replication Machine




The clamp associated with the core polymerase on the lagging strand dissociates at the end of each Okazaki fragment and reassembles for the next fragment.

While one pol synthesizes leading strand the other cycles within the single-stranded template loop (generated by spooling of leading strand) to synthesize the next fragment.

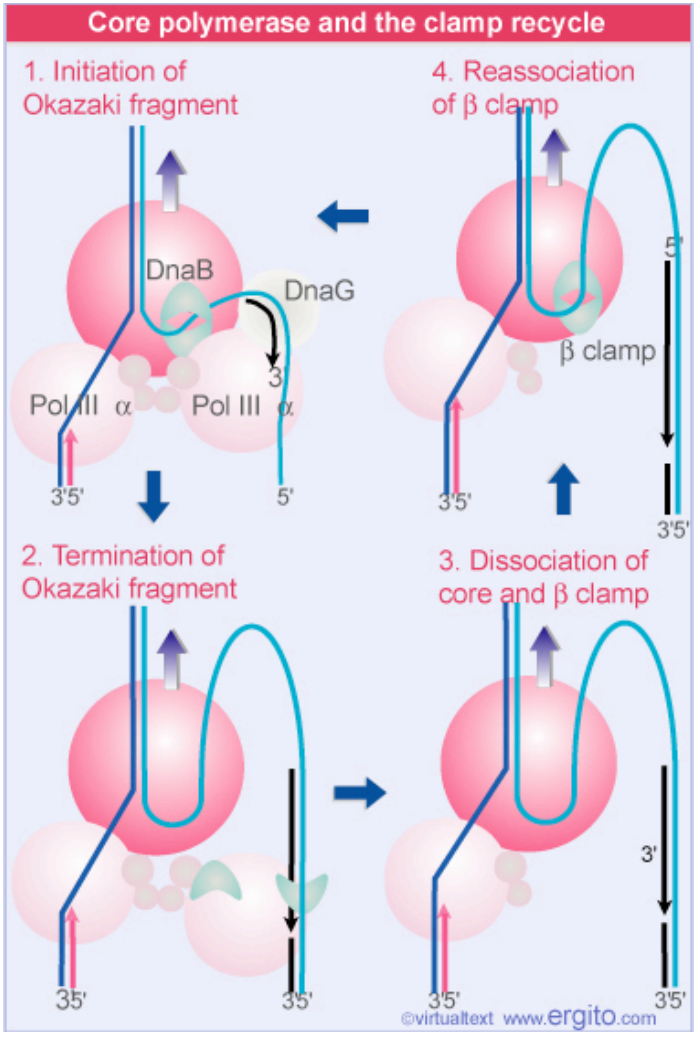
# *E. Coli* Replicase




The helicase DnaB is responsible for interacting with the primase DnaG to initiate each Okazaki fragment

# Lagging Strand Synthesis: Cycles of Loading and Unloading DNA Pol and Clamp Proteins

Helicase DnaB interacts with primase DnaG to signal initiation of Okazaki fragment

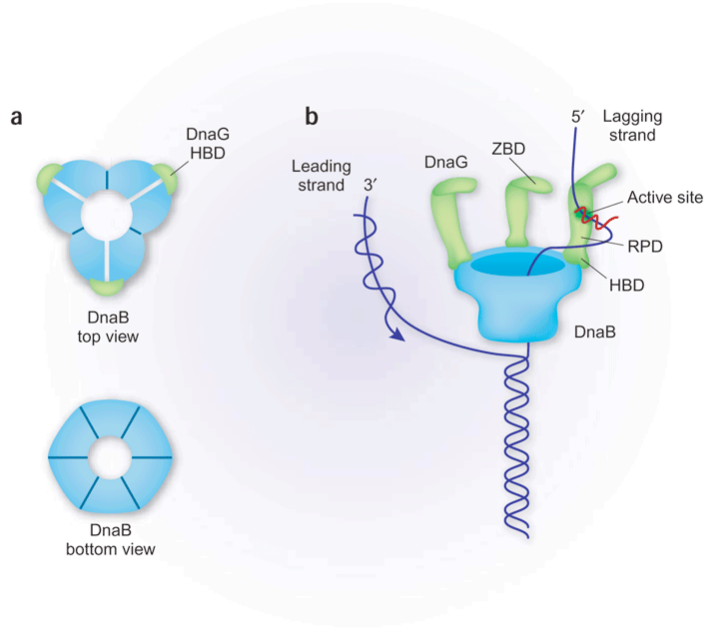



Pol III stays associated with replicating complex

Meets next Okazaki fragment, falls off

DNA Pol is released and associates with a new clamp assembled on the RNA primer of the next Okazaki fragment

# General architecture of the helicase-primase complex; coordination




The lagging strand of the unwound DNA threads through the helicase and is captured by the RNA Pol Domain of one of the associated DnaG molecules; the Zinc binding domain from a different DnaG might participate in this process. The primer is extruded towards the outside of the complex.
HBD= helicase binding domain

K. Marians. Nature Structural & Molecular Biology (2008)

# Replication at the Fork


This animation was built as a "hybrid" of prokaryote and eukaryote replication proteins and dynamics.

eukaryote system speed is 20-50 bases/second with Okazaki fragments 100-150 bases long. The bacterial (prokaryotic) replisome works about 10 times faster, with Okazaki fragments 700-1000 bases long and the incoming DNA rotating at 10,000RPM

two tau subunits (light blue) connect the DNA helicase (dark blue) to the clamp loader (grey fingers), and each tau connects to one of the two polymerase (purple) enzymes.

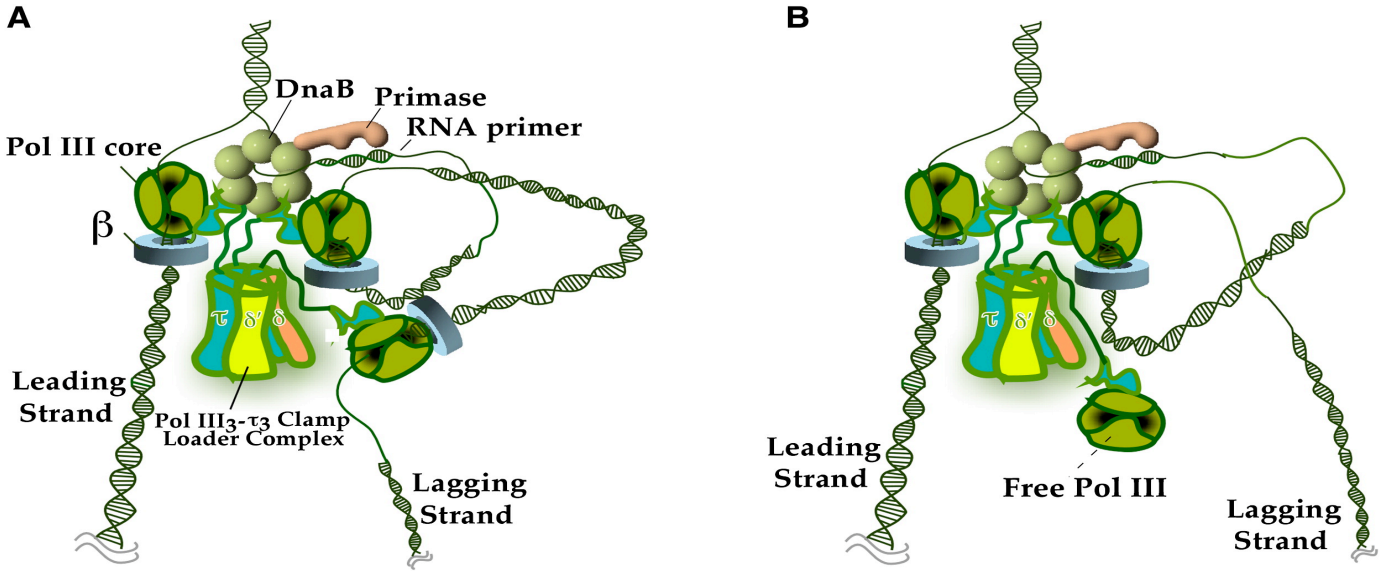
sliding clamps are green and the DNA primase, which is yellow-green, adds RNA primers (yellow) to the ssDNA as it emerges from the helicase.

http://www.wehi.edu.au/education/wehi-tv/dna/replication.html

Note single strand DNA (ssDNA) binding proteins that cap the exposed bases were left out for clarity

Hoatlin Fall 2009

# Rewrite the Textbooks:
# Triple Polymerase Replisome




McInerney at al. Mol Cell 2007. E. coli pol III assembles into a particle that contains THREE DNA polymerases, capable of simultaneous activity. 3 pol is the dominant form.

Two polymerases on the lagging strand, consistent with T4 replication forks (Nossal et al., 2007)

Propose: The third polymerase can act as a reserve enzyme to overcome certain types of replication obstacles or to coordinate the slower lagging strand synthesis

# Emerging Concepts

- Textbook view: single DNA polymerases are bound continuously to their substrates
- Revised view: Polymerases and repair factors are cycled and switched as needed during replication
- The replisome is not dismantled in this process, transient dissociations occur
- Some repair processes previously thought to be "postreplicative" may occur during replication when the replisome is paused

Lovett, Mol Cell 2007

table 25–4

**Proteins at the *E. coli* Replication Fork**

| Protein | $M_r$ | Number of subunits | Function |
|---|---|---|---|
| SSB | 75,600 | 4 | Binding to single-stranded DNA |
| DnaB protein (helicase) | 300,000 | 6 | DNA unwinding; primosome constituent |
| Primase (DnaG protein) | 60,000 | 1 | RNA primer synthesis; primosome constituent |
| DNA polymerase III | 900,000 | 18–20 | New strand elongation |
| DNA polymerase I | 103,000 | 1 | Filling of gaps, excision of primers |
| DNA ligase | 74,000 | 1 | Ligation |
| DNA gyrase (DNA topoisomerase II) | 400,000 | 4 | Supercoiling |

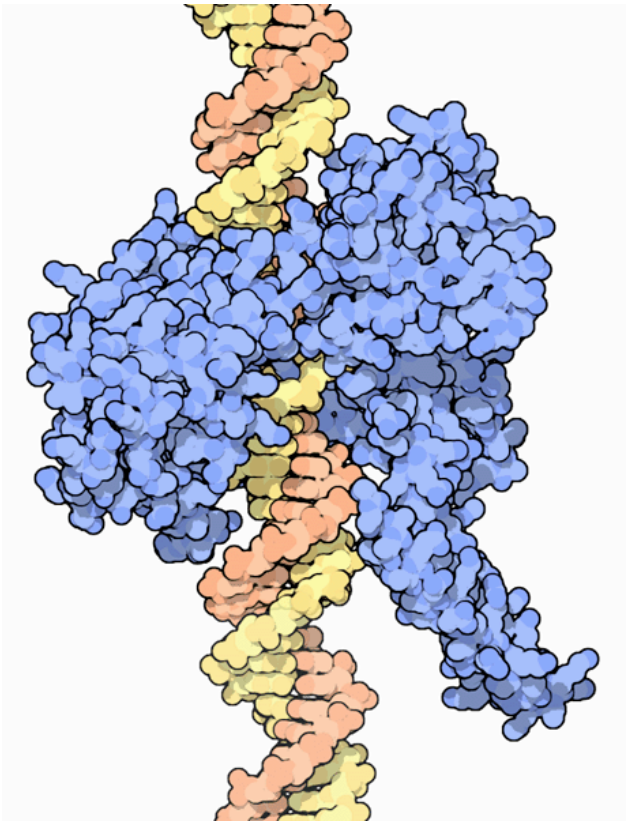
Modified from Kornberg, A. (1982) *Supplement to DNA Replication,* Table S11–2, W.H. Freeman and Company, New York.

# Topoisomerase Classes

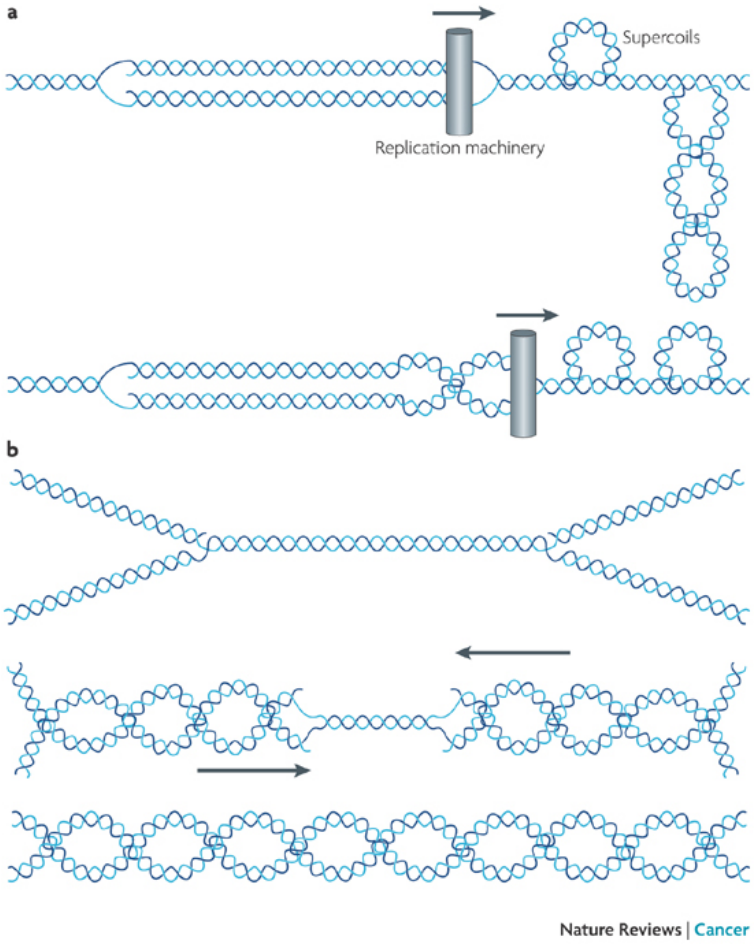
- Topoisomerase I: relaxes DNA
  - -Transient break in one strand of duplex DNA
  - *E. coli*: nicking-closing enzyme
- Topoisomerase II: introduces negative superhelical turns
  - -Breaks both strands of the DNA and passes another part of the duplex DNA through the break; then reseals the break.
  - -Uses energy of ATP hydrolysis
  - -Acts as the SWIVEL for DNA replication.

# Topoisomerases

- Topoisomerases relieve winding/ unwinding tension of DNA during replication fork movement
- Topos wrap around DNA, make a cut which permits the helix to spin and relax, then reconnects strands
- Several drugs interfere with the topoisomerase function. Examples:
  - Fluoroquinolone antibiotic interferes with bacterial type II topoisomerases
  - Topoisomerase II (TOP2) is the target of several important classes of anticancer drugs, including the epipodophyllotoxin etoposide and the anthracycline doxorubicin.



# Roles of topoisomerase II (TOP2) in replication




**Precatenane**
A structure related to a catenane (Circles linked as in a chain) that results from the interwinding of DNA strands behind a replication fork. Precatenanes interconvert with positive supercoils that arise in front of a replication fork.

Helicase creates positive superhelical stress on the DNA

At early steps in replication, when forks are widely separated either TOP1 or TOP2 can function as a replication swivel. TOP1 acts by relaxing positive supercoils whereas TOP2 unlinks precatenanes.

As the replication forks converge, there is a limited ability to generate positive supercoiling, and complete unlinking absolutely requires TOP2

Nitiss, Nature Reviews Cancer, May 2009

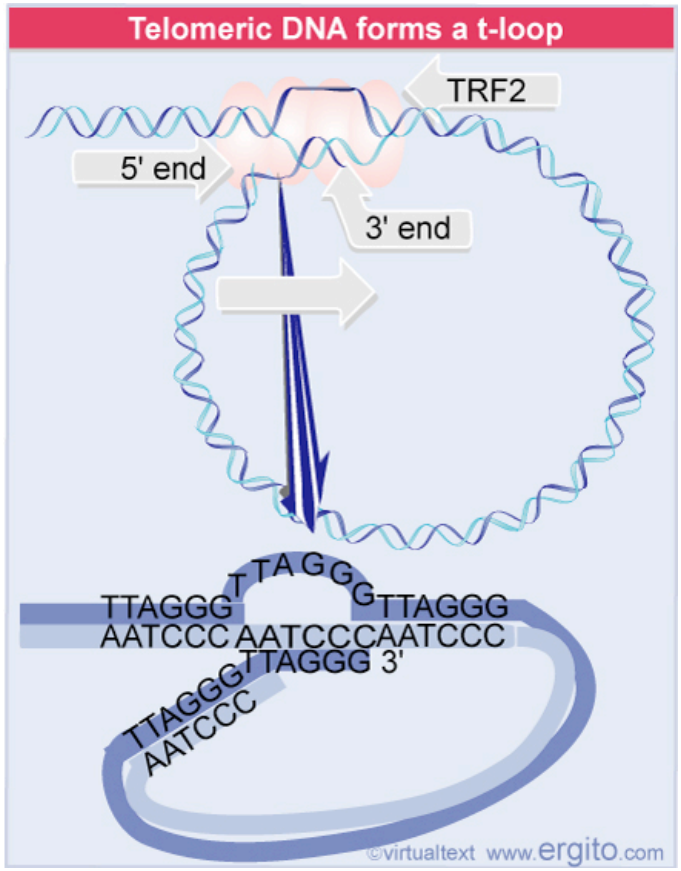
# Winners: Carol Greider, Elizabeth Blackburn and Jack Szostak.

Front

Back

# Telomeres

- Telomeres; series of short, tandemly repeated sequences
- G-T-rich strand that extends beyond a C-A-rich strand
- A loop of DNA forms at the telomere, sealing the chromosome ends and stabilizing the chromosome
- Extension of the ends of the replicating linear chromosome is performed by **telomerase,** a large ribonucleoprotein enzyme that provides the template for extension of the C-A-rich strand




Hoatlin Fall 2009

- "...there is still a lot of basic biology to discover — such as how telomerase activity is regulated at individual telomeres, and how telomeres manage to avoid the attentions of DNA repair enzymes which seek out breaks in DNA and restitch the torn ends. "

# Overview of Replication

- Of the 165 genes that are involved in replicating the human genome only 3 are not found in yeast
- The sequence of events that initiate, elongate and terminate DNA replication is essentially the same throughout the eukaryotic world
- >15 proteins are required to assemble eukaryotic pre-replication complexes
- >20 additional proteins are required to assemble replication forks and initiate DNA synthesis
- Many more proteins are required to maintain forks, assemble newly replicated DNA into chromatin, methylate DNA, respond to DNA damage and maintain the ends of chromosomes

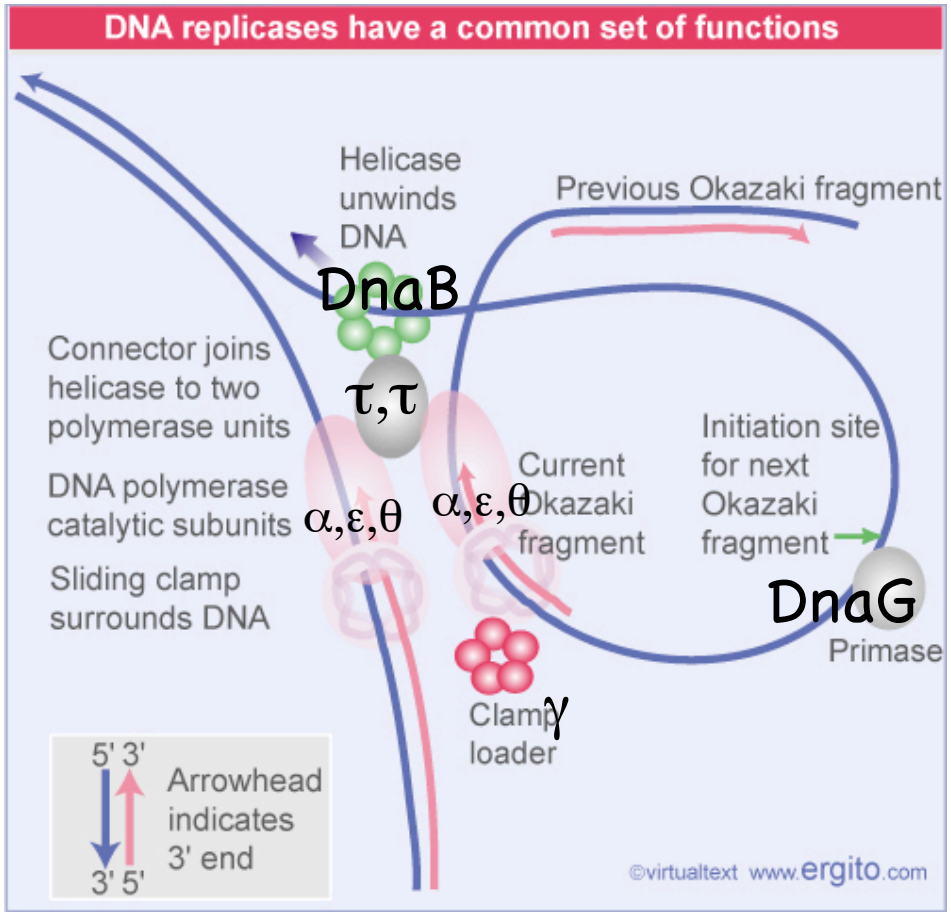
Hoatlin Fall 2009

# DNA Replication II

- Eukaryotic versions of polymerases, clamps and clamp loaders
- The prokaryotic system
  - Initiation: begins at a specific site, e.g. *oriC* for *E. coli*.
  - Elongation: movement of the replication fork
  - Termination: at *ter* sites for *E. coli*
- Eukaryotic replication initiation
- Helicases: replication and other functions
- What next?

# Eukaryotic Replication

# The Replication Machine




Review:
DnaG: primase
[α,ε,θ] DNA Pol III catalytic subunits
β- sliding clamp
γ-clamp loader
DnaB: helicase

Hoatlin Fall 2009

# Eukaryotic cells have many DNA polymerases

replication

repair

| Eukaryotic DNA polymerases undertake either replication or repair | | | |
|---|---|---|---|
| DNA polymerase | | Function | Structure |
| | | High fidelity replicases | |
| α | alpha | Nuclear replication | 350 kD tetramer |
| δ | delta | " | 250 kD tetramer |
| ε | epsilon | " | 350 kD tetramer |
| γ | gamma | Mitochondrial replication | 200 kD dimer |
| | | High fidelity repair | |
| β | beta | Base excision repair | 39 kD monomer |
| | | Low fidelity repair | |
| ζ | zeta | Thymine dimer bypass | heteromer |
| η | eta | Base damage repair | monomer |
| ι | iota | Required in meiosis | monomer |
| κ | kappa | Deletion and base substitution | monomer |

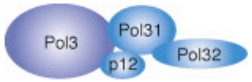


Replicative enzymes operate with high fidelity.

Except for the β enzyme, the repair enzymes all have low fidelity.

Replicative enzymes have large structures, with separate subunits for different activities. Repair enzymes have much simpler structures.

Major DNA polymerases at the replication fork[a]

| | Pol α-primase | Pol δ | Pol ε |
|---|---|---|---|
| Subunit organization |   |   |   |
| **Genes and subunit sizes** | | | |
| *S. cerevisiae* | Pol1-p167 | Pol3-p125 | Pol2-p256 |
| | Pol12-p79 | Pol31-p55 | Dpb2-p78 |
| | Pri1-p48 | Pol32-p40 | Dpb3-p23 |
| | Pri2-p62 | – | Dpb4-p22 |
| *S. pombe* | Pol1-p159 | Pol3-p124 | Pol2-p253 |
| | Pol12-p64 | Cdc1-p51 | Dpb2-p67 |
| | Pri1-p52 | Cdc27-p42 | Dpb3-p22 |
| | Spp2-p53 | Cdm1-p19 | Dpb4-p24 |
| Human | PolA1-p166 | PolD1-p124 | PolE-p261 |
| | PolA2-p68 | PolD2-p51 | PolE2-p59 |
| | Prim1-p48 | PolD3-p66 | PolE3-p17 |
| | Prim2A-p58 | PolD4-p12 | PolE4-p12 |
| Activity | Polymerase | Polymerase | Polymerase |
| | Primase | 3′-exonuclease | 3′-exonuclease double-strand-DNA binding |
| Fidelity | $10^{-4}$–$10^{-5}$ | $10^{-6}$–$10^{-7}$ | $10^{-6}$–$10^{-7}$ |
| Function | Initiation of replication | Elongation and maturation of Okazaki fragments | Replisome assembly |
| | Initiation of Okazaki fragments | DNA repair | Leading-strand synthesis |
| | | Mutagenesis | Replication checkpoint |

[a] The nomenclature for the cartoon depictions is for S. cerevisiae genes. For Pol δ, a fourth subunit (p12) is shown, which is found in humans but not

# Eukaryotic Replicative Pols

- DNA pol $\alpha$- initiates new strands leading and lagging (unusual! called pol $\alpha$/primase)
  - limited processivity, lacks intrinsic 3′ exonuclease activity for proofreading errors, it is not well suited to efficiently and accurately copy long templates.
  - initiates replication at origins and during lagging-strand synthesis of Okazaki fragments
- DNA pols $\varepsilon$ , $\delta$ - Elongation

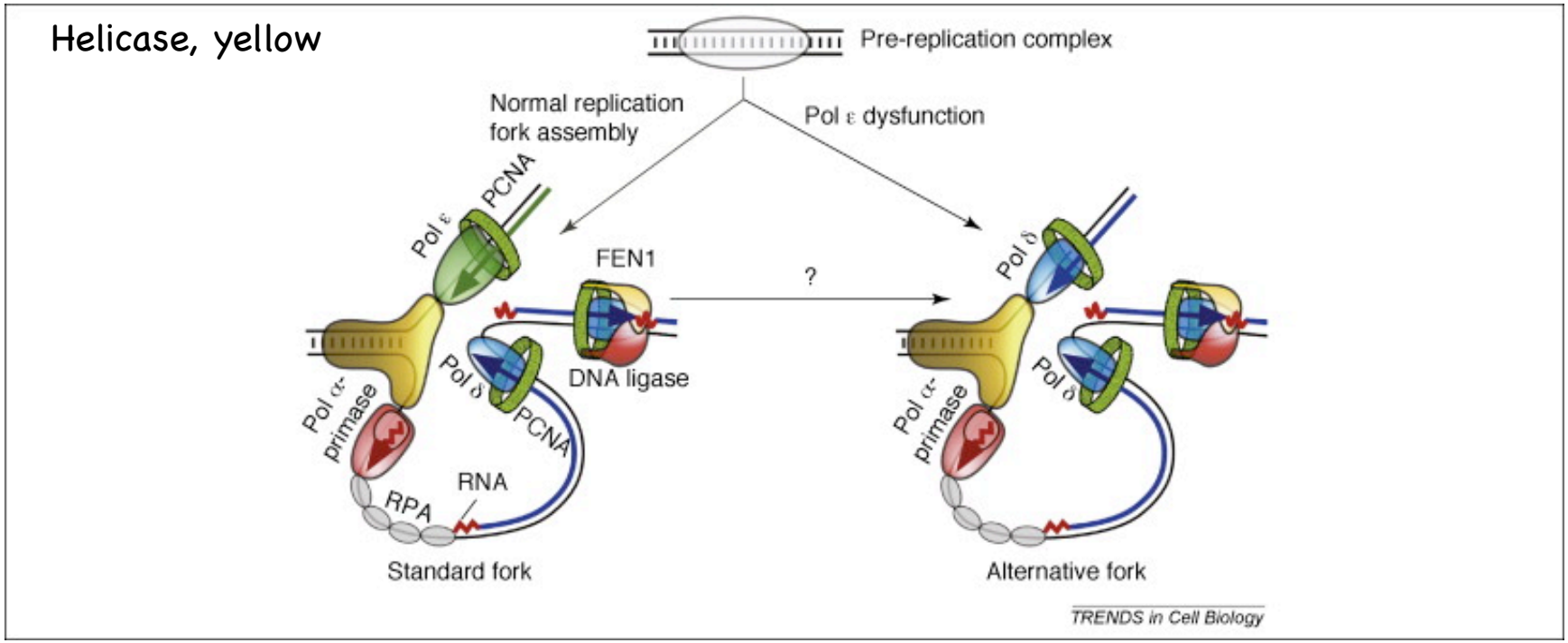
Endonucleases are enzymes that cleave the phosphodiester bond within a polynucleotide chain, in contrast to exonucleases, which cleave phosphodiester bonds at the end of a polynucleotide chain.

# Pol ε and pol δ

- Intrinsic 3′ exonucleolytic proofreading activities
- DNA pol ε primarily involved in leading strand synthesis (N.B. most textbooks and many ref. papers may be incorrect on this point)
  - Kunkel Lab: Yeast DNA polymerase epsilon participates in leading-strand DNA replication. Science 2007. Created defective polymerase epsilon with a signature reduced fidelity to mark leading strand. Landmark advance.
- DNA pol δ- primarily involved in lagging strand synthesis

# Models for eukaryotic DNA replication forks

Helicase, yellow




*the flap endonuclease FEN1 degrades the initiator RNA during Okazaki-fragment maturation.

Kunkel & Burgers. Trends in Cell Biology, v 18, Nov 2008!!!

# Eukaryotic Replication

- DNA polymerase $\alpha$/primase binds an initiation complex at the origin and synthesizes 10 bases of RNA followed by 20-30 bases of DNA (iDNA)
- RFC clamp loader binds to 3′ end of iDNA and and uses ATP hydrolysis to open the ring of PCNA (PCNA is like the $\beta$ sliding clamp in structure but no sequence similarity)--see next slide
- Pol Switch- $\varepsilon$ on leading interacts with RFC (clamp loader) and PCNA (sliding clamp)
- Lagging strand-elongation by pol $\delta$
- There are many pols...

# Human DNA Polymerases

| Polymerase[a] | Family | Catalytic subunit | | | | Associated activities | Proposed functions |
|---|---|---|---|---|---|---|---|
| | | Molecular mass (kDa)[b] | Human gene (alias) | Chromosomal location[c] | Yeast gene[d] (alias) | | |
| α (alpha) | B | 165 | *POLA* | Xp22.1-p 21.3 | *POL1 (CDC17)* | Primase | chromosomal replication, S-phase checkpoint, DSB repair |
| β (beta) | X | 39 | *POLB* | 8p11.2 | - | dRP & AP lyase | BER, single strand break repair |
| γ (gamma) | A | 140 | *POLG* | 15q25 | *MIP1* | 3'→5' exonuclease, dRP lyase | mitochondrial replication, mitochondrial BER |
| δ (delta) | B | 125 | *POLD1* | 19q13.3 | *POL3 (CDC2)* | 3'→5' exonuclease | chromosomal replication, NER, BER, MMR, DSB repair |
| ε (epsilon) | B | 255 | *POLE* | 12q24.3 | *POL2* | 3'→5' exonuclease | chromosomal replication. NER. BER, MMR, DSB repair, S-phase checkpoint |
| ζ (zeta) | B | 353 | *POLZ (REV3)* | 6q21 | *REV3* | | TLS. DSB repair. ICL repair?. SHM |
| η (eta) | Y | 78 | *POLH (RAD30, RAD30A, XPV)* | 6p21.1 | *RAD30* | | TLS, SHM |
| θ (theta) | A | 198 | *POLQ* | 3q13.33 | - | | ICL repair? |
| ι (iota) | Y | 80 | *POLI (RAD30B)* | 18q21.1 | - | dRP lyase | TLS?, BER?, SHM |
| κ (kappa) | Y | 76 | *POLK (DINB1)* | 5q13 | - | | TLS |
| λ (lambda) | X | 66 | *POLL* | 10q23 | *POL4 (POLX)* | dRP lyase | DSB repair, BER? |
| μ (mu) | X | 55 | *POLM* | 7p13 | - | TdT | DSB repair |
| σ (sigma) | X | 60 | *POLS (TRF4-1)* | 5p15 | *TRF4* | | sister chromatid cohesion |
| REV1 | Y | 138 | *REV1* | 2q11.1-q11.2 | *REV1* | TdT (for dC) | TLS |

Shcherbakova, et al. 2003

# Special Function Eukaryotic Polymerases

- Replicative bypass of DNA damage or in specialized repair that would block normal replication
- Signal that the cell is not ready to progress in the cell cycle towards mitosis

# Pols with Special Powers

- Additional enzymatic activities
- Pol I ι (iota) misinserts a G opposite a template T much better than the correct A.
- Implicated in repair of interstrand crosslinks: Pol ζ (*REV3*), Pol θ (*mus308*)
- Some Pols have "hand" with a snug fit and cannot accommodate bulky lesions➜ replication block, while Pol ζ, and the Y family can bypass (TLS or translesion synthesis)

# The Y Family (η, ι, κ, Rev1)

- Extra DNA binding domain, and more open "hand"
- Very low fidelity reflecting lower selectivity
- Factory model: many pols are present at the replication complex, called on to function when replicative pols are stalled, and then allow main pols to resume
- Note: defective Pol η (eta) in Xeroderma Pigmentosum, defect in removing cyclobutane pyrimidine dimers ➜➜ skin cancer

# Eukaryotic Clamps and Clamp Loaders: similar theme in the three domains of life

| *Replicase Component* | *E. coli* | *Eukaryotes* | *Archaea* | *T4 Phage* |
|---|---|---|---|---|
| **Polymerase** | Core (αεθ) | Pol δ/Pol ε | Pol δ | gp 43 |
| **Sliding Clamp** | β clamp | PCNA | PCNA | gp 45 |
| **Clamp Loader** | γ complex | RFC | RFC | gp 45/62 |

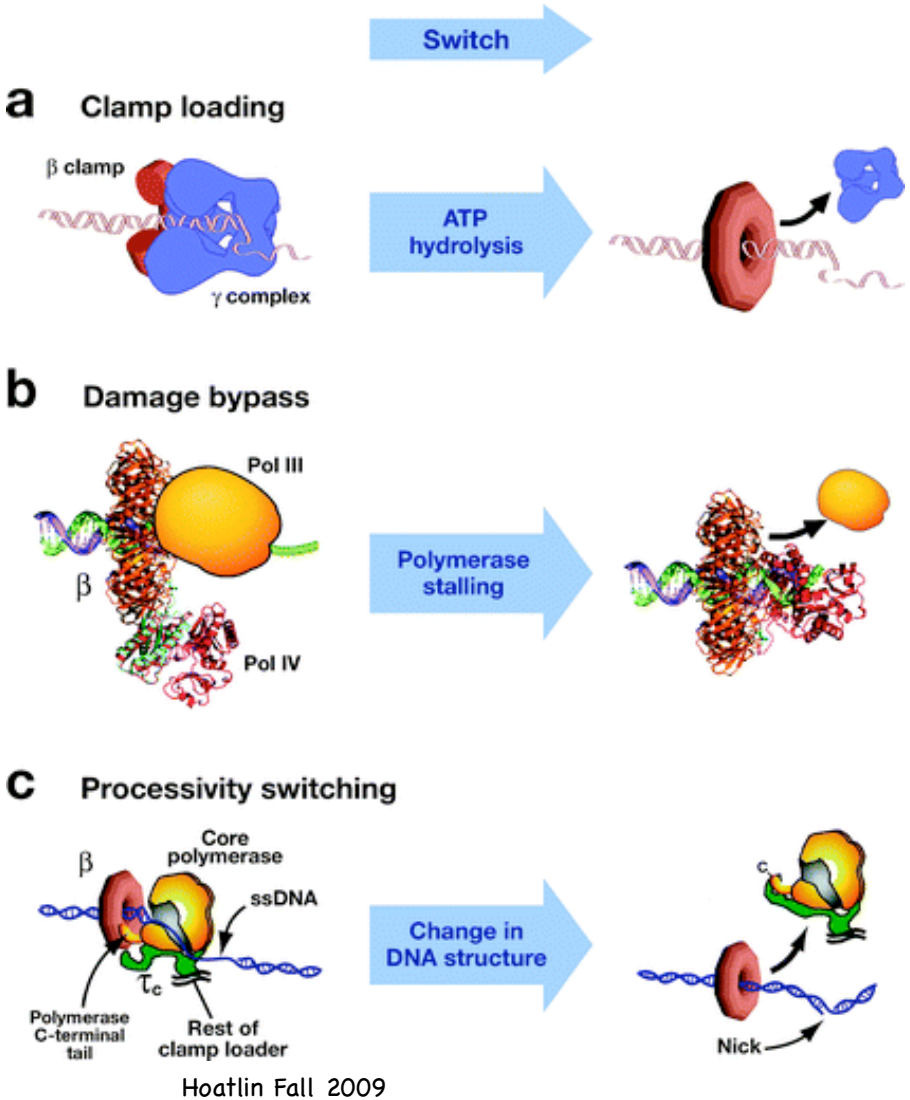
# The Sliding Clamp--Directing Traffic for the Pol Switch in TLS?

PCNA (clamp) becomes modified by monoubiquitination after DNA damage

PCNA-Ub, but not unmodified PCNA, interacts with DNA polymerase eta

monoubiquitination of Pol $\eta$ (eta) may be central to polymerase switching during TLS (Kannouche PL, Wing J, Lehmann AR. *Mol Cell* 2004 14(4):491-500)

# Examples of protein trafficking on sliding clamps



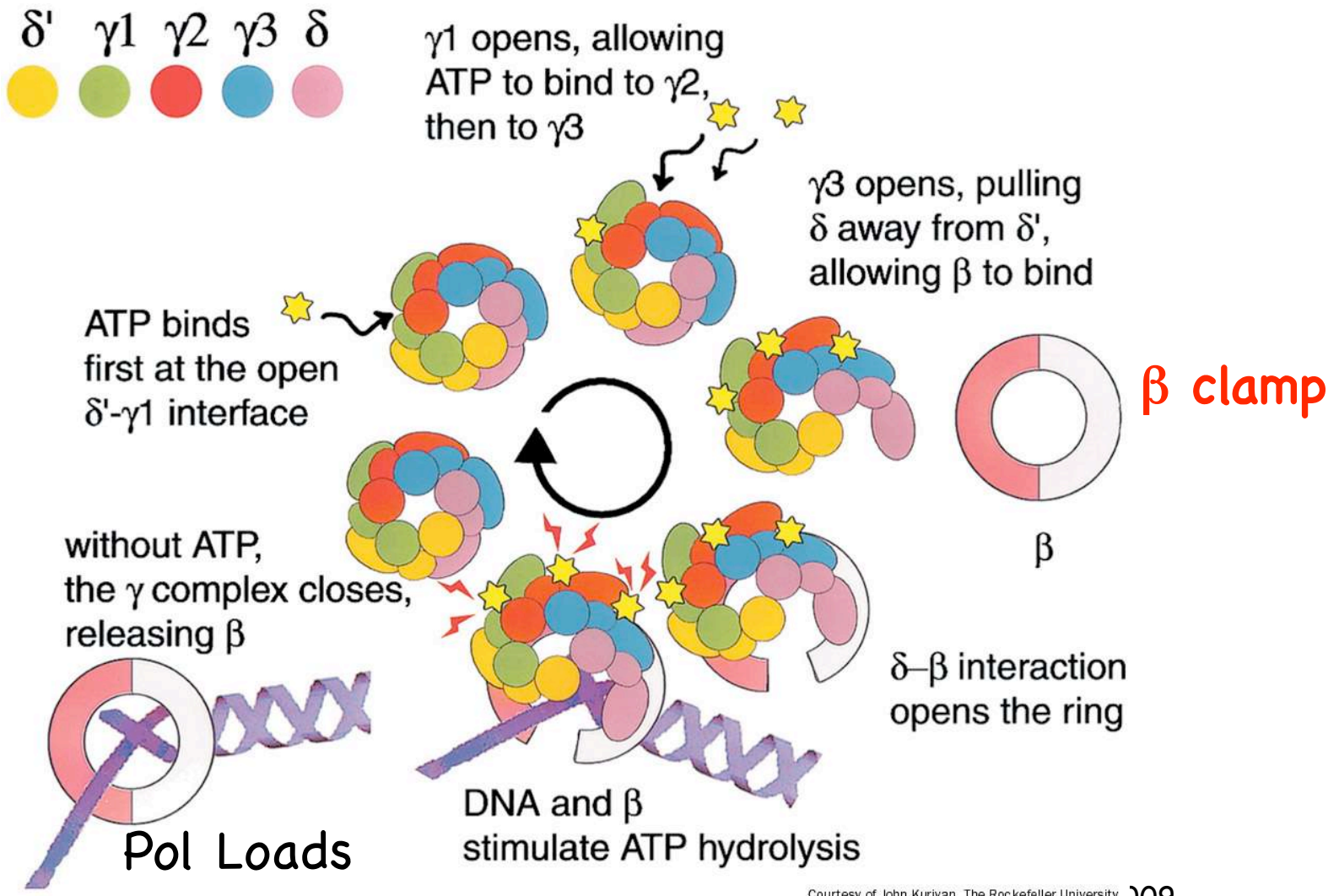

Hoatlin Fall 2009

The $\gamma$ complex clamp loader associates tightly with $\beta$ when bound to ATP. DNA triggers ATP hydrolysis, resulting in low affinity for $\beta$ and DNA

When Pol III encounters a lesion in the DNA template, it stalls, unable to overcome its inherent fidelity to incorporate opposite a damaged base. Stalling allows an error-prone polymerase, such as Pol IV (red) passively traveling on $\beta$, an opportunity to trade places with Pol III on $\beta$ to replicate past the lesion.

Pol III has a tight grip on $\beta$ until complete replication of its substrate DNA, when the polymerase releases from $\beta$ to recycle to the next primed site. The $\tau$ s.u. modulates this interaction, binding the polymerase C tail only when no more single-stranded template is present. This severs the connection between the polymerase and the clamp.

# Recall the Clamp Loader Cycle (*E. coli*)



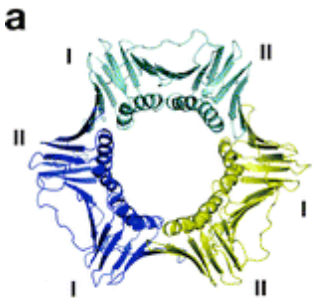

Courtesy of John Kuriyan, The Rockefeller University

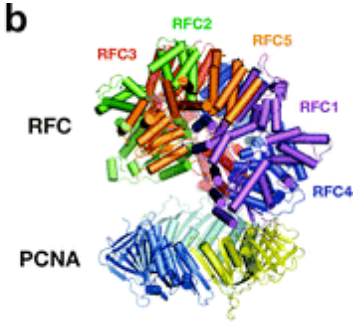
# Eukaryotic Clamp Loading

- New structural studies: Miyata et al, Nature Structural Biology (archaeal), 11:632. 2004 and Bowman et al, Nature, 429:724, 2004 (yeast)
- 5 subunit assembly, ($\gamma$ equivalent clamp loader) Replication Factor C (RFC) with PCNA (clamp) and DNA helix using a hydrolytically inert ATP$\gamma$S
- Structure suggests a"screw cap" assembly, with the pitch of the RFC subunits matching the pitch of the B form of helical DNA--replacement of the clamp loader with the polymerase on PCNA would position the pol for DNA synthesis

# Structures of the eukaryotic clamp and clamp loader from S. cerevisiae

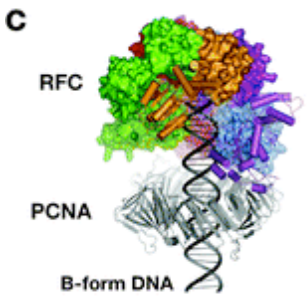
C-terminal face of PCNA

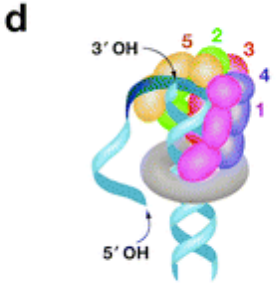


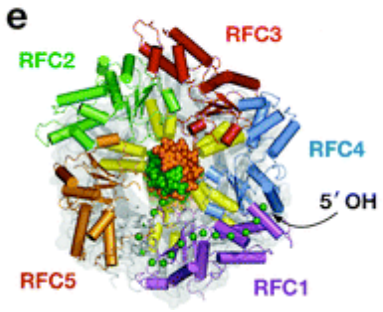

Replication factor C (RFC) bound to PCNA reveals the structural similarity between RFC and $\gamma$ complex

The RFC subunits are arranged in a helix that tracks the minor groove of B-form DNA modeled through the PCNA ring
*new structural studies




5' terminus of a recessed primer template is positioned to exit the central channel of the clamp and clamp loader through the gap between RFC1 and RFC5




Two conserved helices in each RFC subunit (yellow) are in position to interact with DNA (orange/green) that passes through the central channel of PCNA (gray) with the 5' terminus (green spheres) exiting between RFC1 and RFC5

Bowman GD, O'Donnell M, Kuriyan J. 2004. Nature

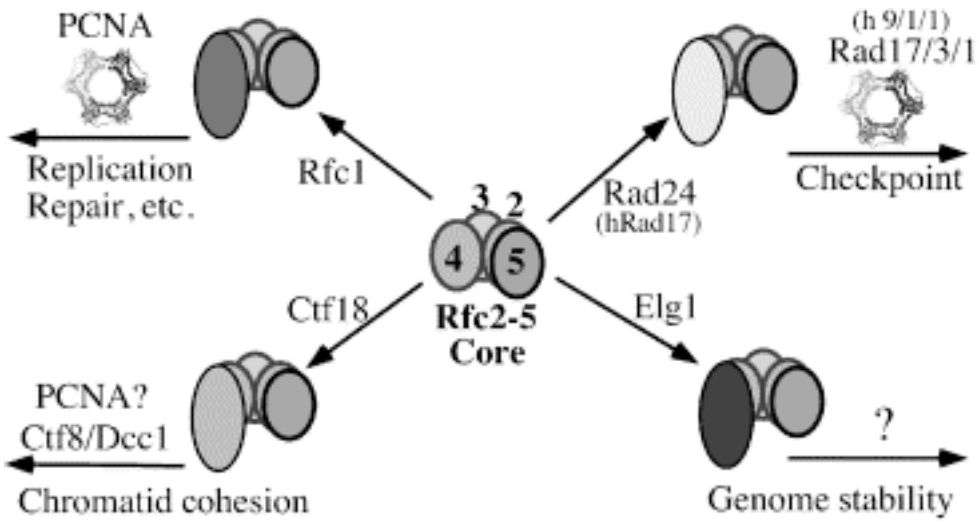
Hoatlin Fall 2009

# Alternative Clamp Loaders

- In response to DNA damage, RFC1 is replaced by a Rad protein (Rad24 in S. cerevisiae, Rad17 in humans);DNA Repair. 2004.

- Rad-RFC loads the clamp (the 911 clamp, which is different from PCNA) onto the 5' end of a primed site and thus recognizes the opposite side of a single-stranded/double-stranded junction compared to RFC-PCNA.

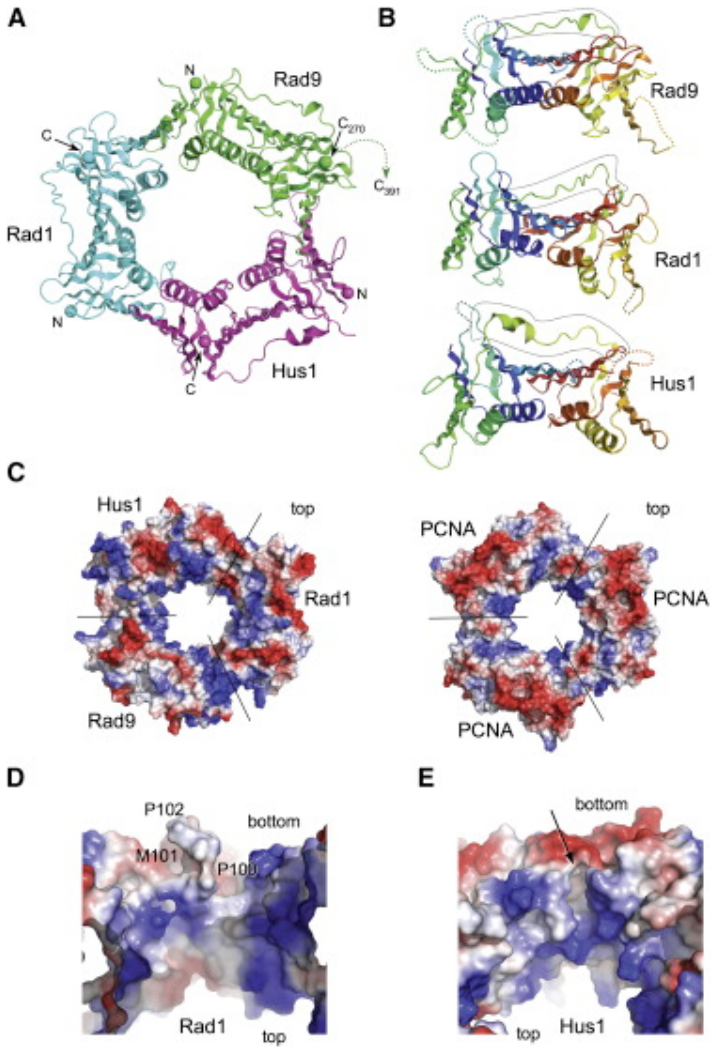
# Alternative clamp-clamp loader systems




All these clamp loaders have the Rfc2-5 core in common with RFC. However, Rfc1 is replaced by a pathway-specific Rfc1-like protein, Rad24 for the DNA damage checkpoint, Ctf18 for the establishment of chromatid cohesion, and Elg1 for an ill-defined pathway that functions in the maintenance of chromosome stability.

From Majka & Burgers. Progress in Nucleic Acid Research and Molecular Biology 2004.

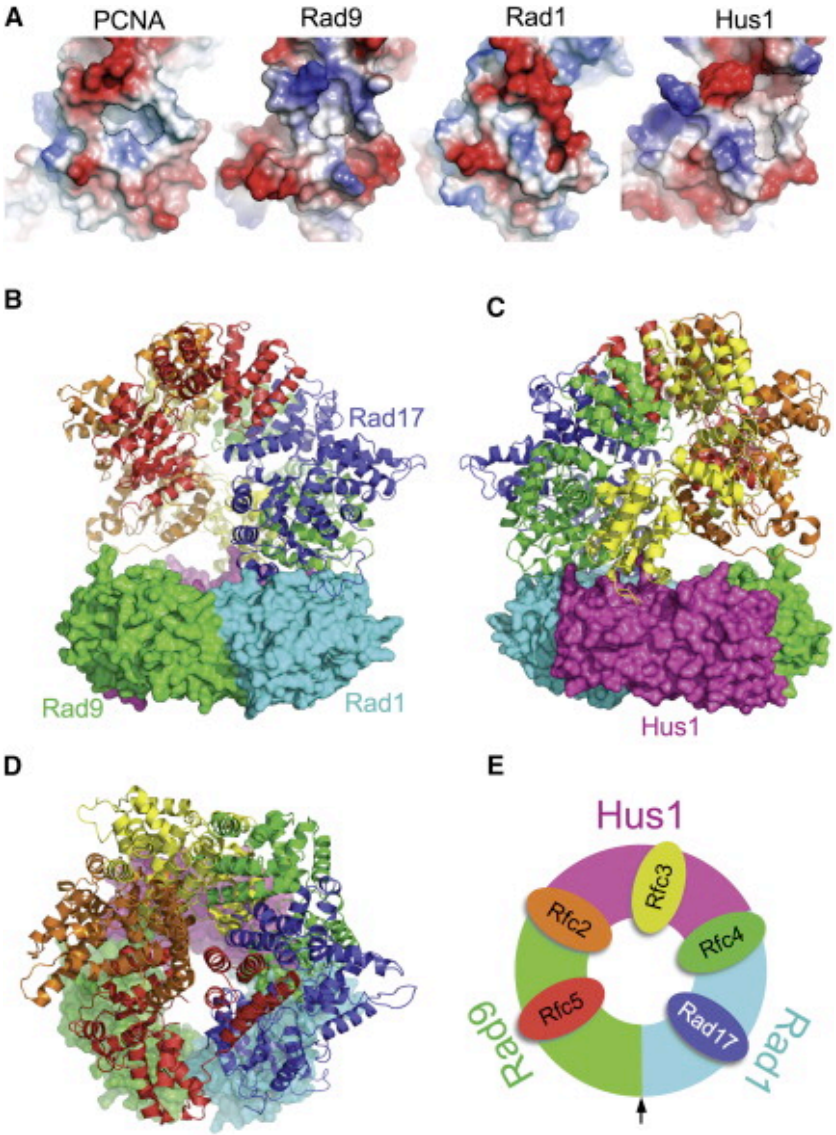
# Structure of the Rad9-Rad1-Hus1 Complex




Dore, et al., Mol Cell 2009

Hoatlin Fall 2009

# Clamp Loader Interactions with 9-1-1




Hoatlin Fall 2009

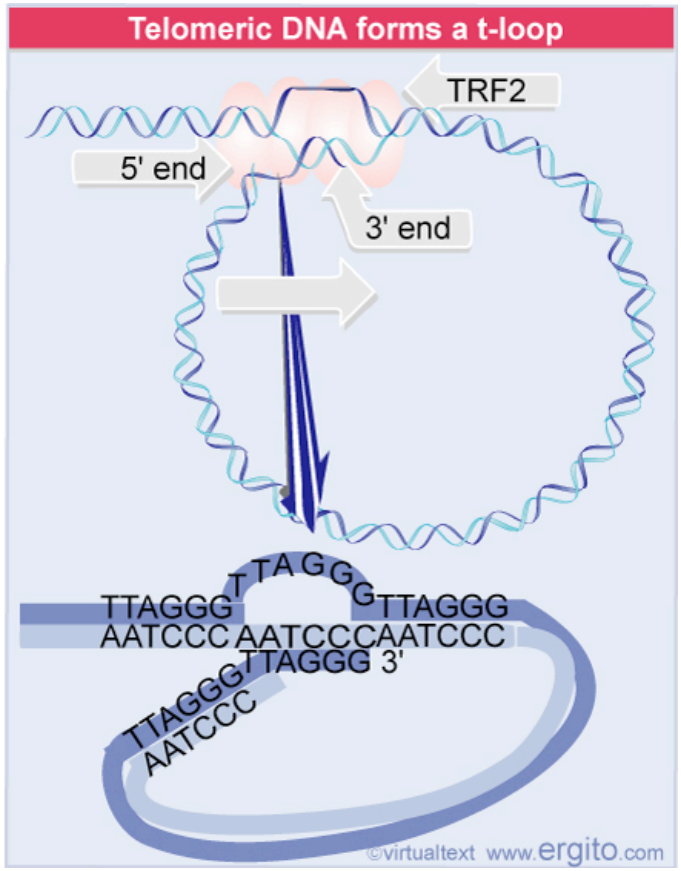
# Questions

- The nature of the ATP-induced conformational change in the clamp loader needed for clamp and DNA binding remains uncertain. Is the clamp also required for this conformational change?
- How does the clamp loader read the polarity of a primer-template junction? Does it recognize the 3′ or 5′ terminus at a primed site? Perhaps instead it reads the polarity of the template strand.
- How are alternative RFC complexes coordinated?
- Both prokaryotic and eukaryotic clamps interact with a wide variety of DNA polymerases, repair factors and cell cycle regulators. How is this coordinated?

From O'Donnell & Kuriyan, Current Opinion in Structural Biology. 2006

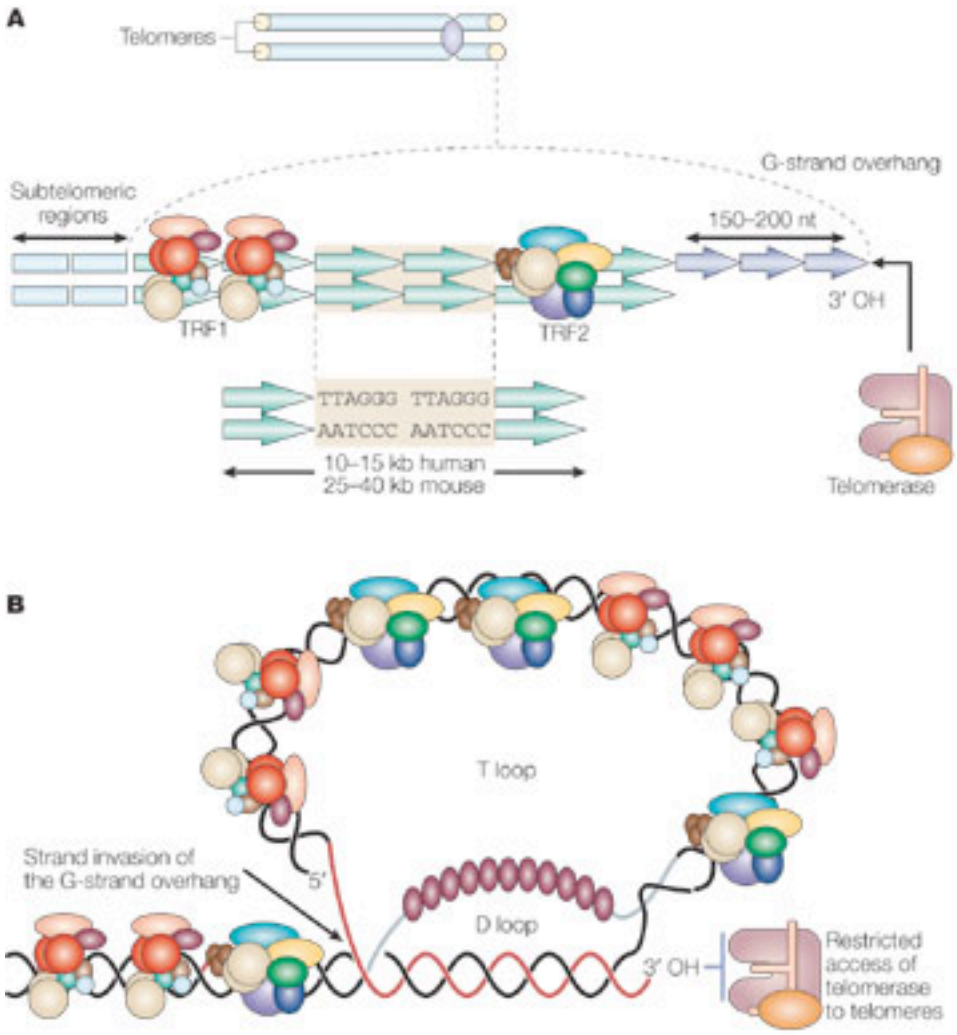
# Telomeres

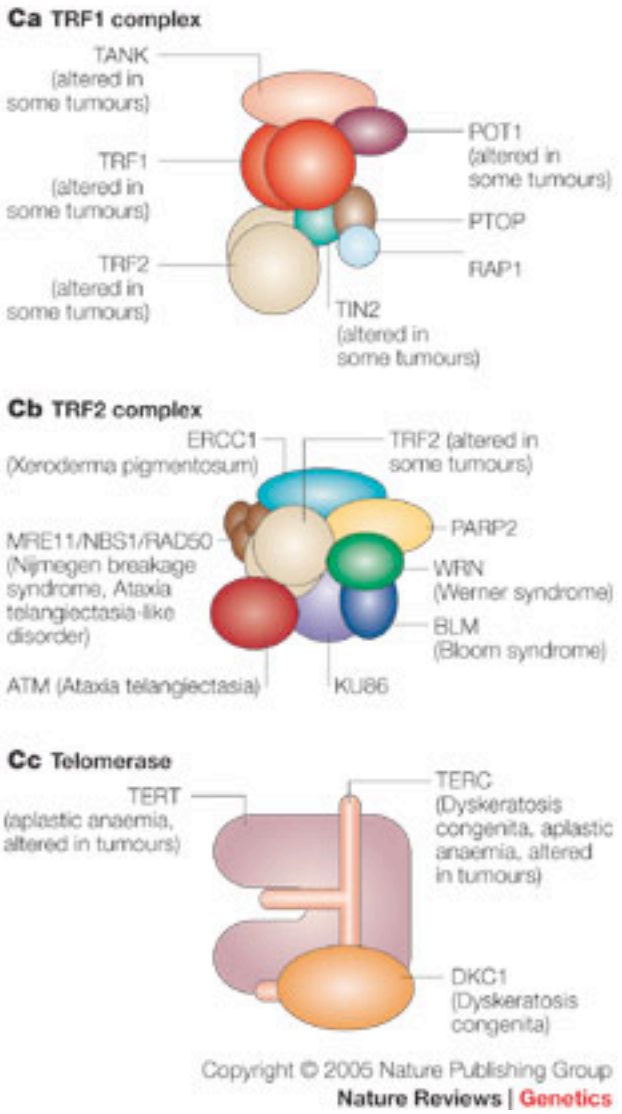
- Telomeres; series of short, tandemly repeated sequences
- G-T-rich strand that extends beyond a C-A-rich strand
- A loop of DNA forms at the telomere, sealing the chromosome ends and stabilizing the chromosome
- Extension of the ends of the replicating linear chromosome is performed by **telomerase**, a large ribonucleoprotein enzyme that provides the template for extension of the C-A-rich strand
- See figs 7.25 & 26 pgs 288-9 in text




Hoatlin Fall 2009

# The structure of mammalian telomeres

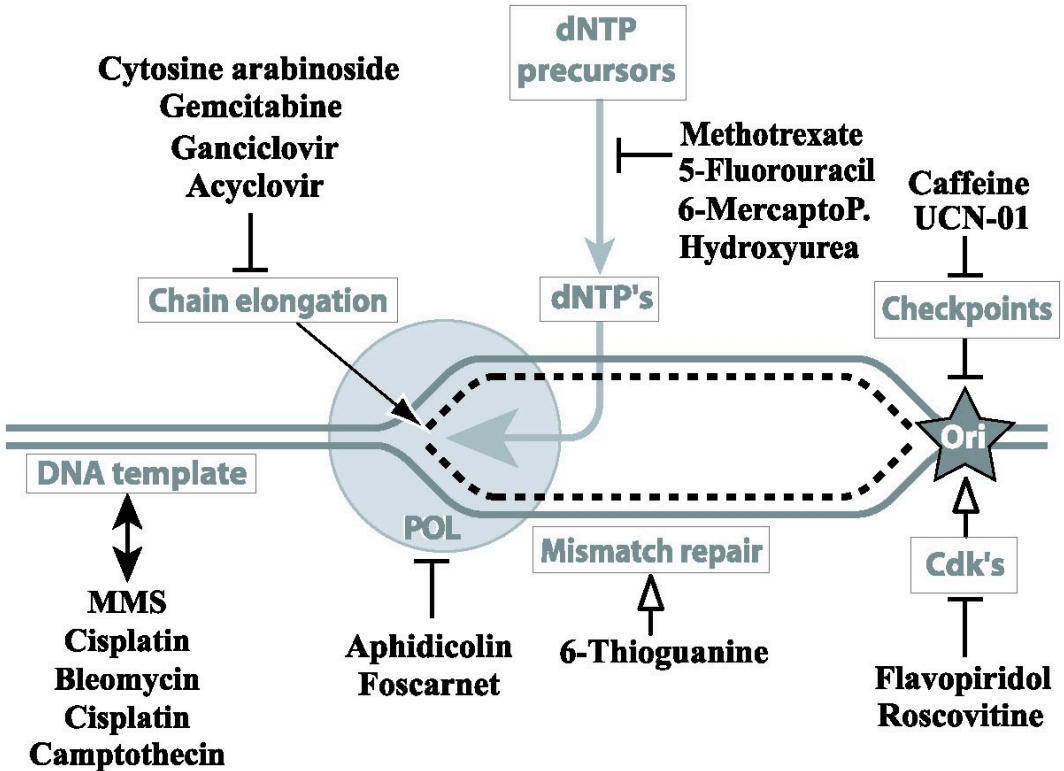





M.Blasco, Nat Rev Genet

# Clinical Significance: Replication



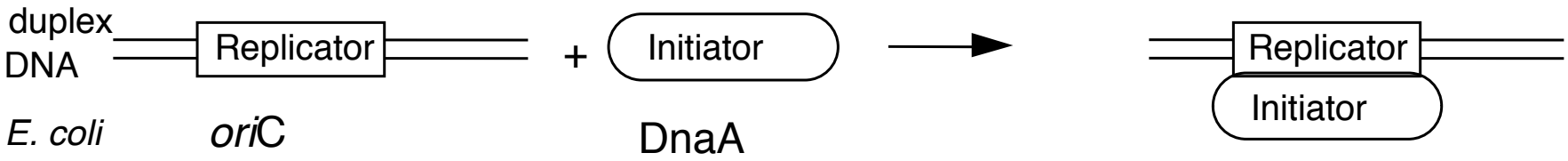

Topo II inhibitor; treatment for lung, ovarian, testicular cancers

Figure 1. Site of action of commonly used DNA replication inhibitors. 6-Mercaptopurine is abbreviated 6-MercaptoP., and 7-hydroxystaurosporine is abbreviated UCN-01.

*DNA Replication and Human Disease* © 2006 Cold Spring Harbor Laboratory Press, Chapter 26, Figure 1.

# Initiation

# Replicon: a unit that controls replication



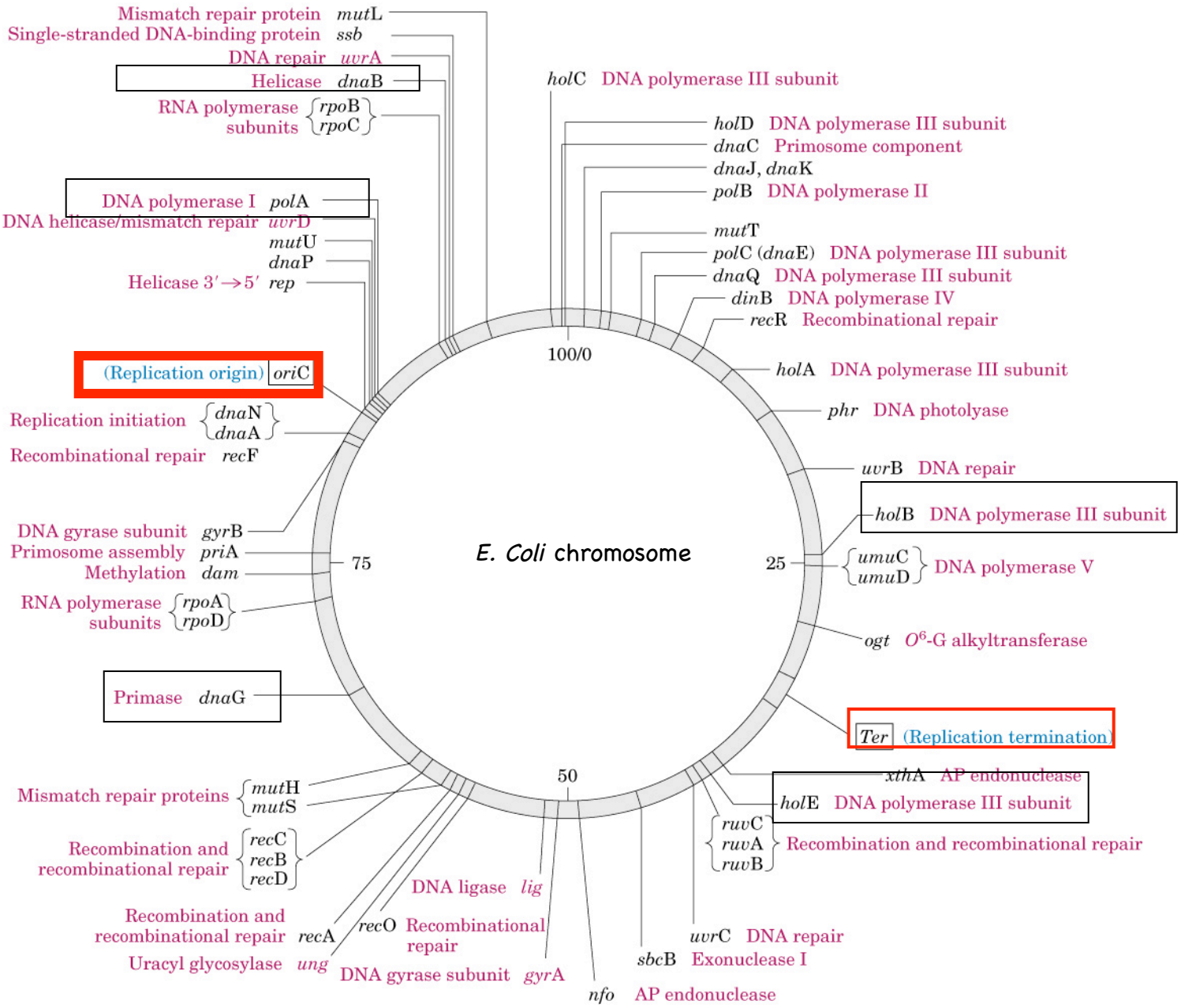

*E. coli* — *oriC* — DnaA

**Yeast** — ***ARS*** **an autonomously replicating sequence** — **ORC (the origin recognition complex) + ABF1 (ARS binding factor 1)** — **Complex of Replicator + initiator allows replication to begin**

Replicator: *cis*-acting DNA sequence required for initiation; defined genetically

Origin: site at which DNA replication initiates; defined biochemically

Initiator: protein needed for initiation, acts in *trans*

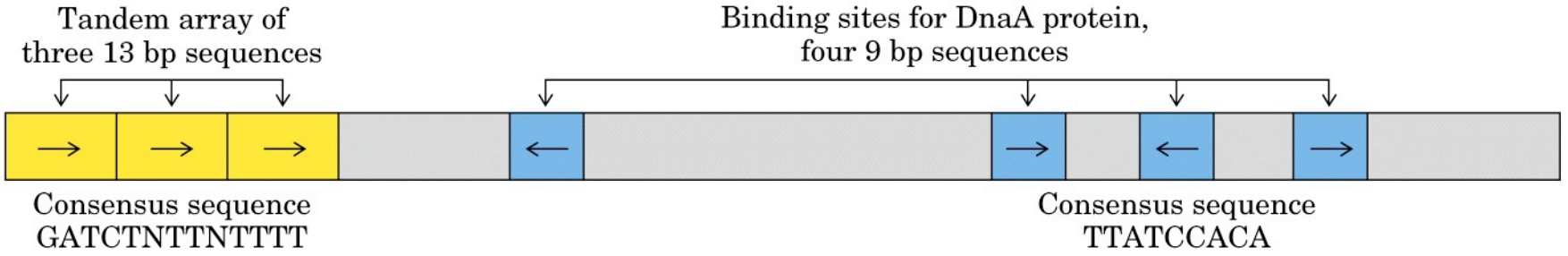

E. Coli chromosome
100/0
25
50
75
Mismatch repair protein mutL
Single-stranded DNA-binding protein ssb
DNA repair uvrA
Helicase dnaB
RNA polymerase subunits rpoB rpoC
DNA polymerase I polA
DNA helicase/mismatch repair uvrD
mutU
dnaP
Helicase 3′→5′ rep
(Replication origin) oriC
Replication initiation dnaN dnaA
Recombinational repair recF
DNA gyrase subunit gyrB
Primosome assembly priA
Methylation dam
RNA polymerase subunits rpoA rpoD
Primase dnaG
Mismatch repair proteins mutH mutS
Recombination and recombinational repair recC recB recD
Recombination and recombinational repair recA
Uracyl glycosylase ung
recO Recombinational repair
DNA ligase lig
DNA gyrase subunit gyrA
nfo AP endonuclease
sbcB Exonuclease I
uvrC DNA repair
ruvC ruvA ruvB Recombination and recombinational repair
holE DNA polymerase III subunit
xthA AP endonuclease
Ter (Replication termination)
ogt O6-G alkyltransferase
umuC umuD DNA polymerase V
holB DNA polymerase III subunit
uvrB DNA repair
phr DNA photolyase
holA DNA polymerase III subunit
recR Recombinational repair
dinB DNA polymerase IV
dnaQ DNA polymerase III subunit
polC (dnaE) DNA polymerase III subunit
mutT
polB DNA polymerase II
dnaJ, dnaK
dnaC Primosome component
holD DNA polymerase III subunit
holC DNA polymerase III subunit

# Common events in priming replication at the origin

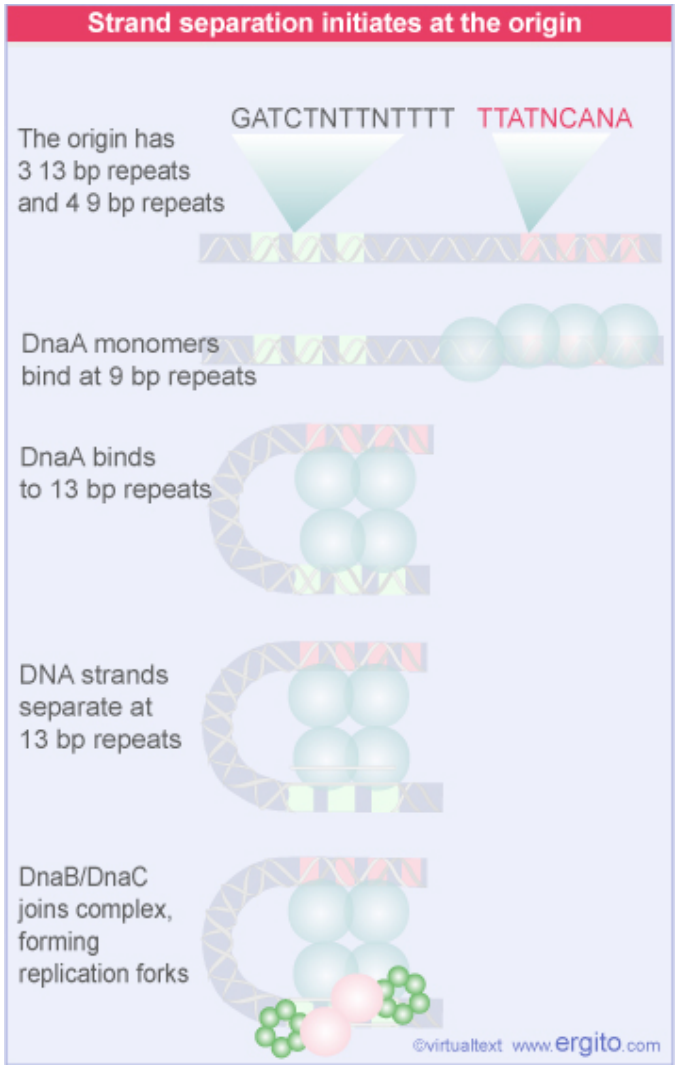
- The general principle of bacterial initiation is that the origin is initially recognized by a protein that forms a large complex with DNA
- A short region of A·T-rich DNA is melted
- DnaB is bound to the complex and creates the replication fork

# Priming in *E. coli*

- E. coli contains the oriC origin of replication
- oriC was identified by its ability to confer autonomous replication on a DNA molecule, thus it is a replicator
- Replication from oriC is bidirectional
- Structure of oriC:



# Creating replication forks at an origin

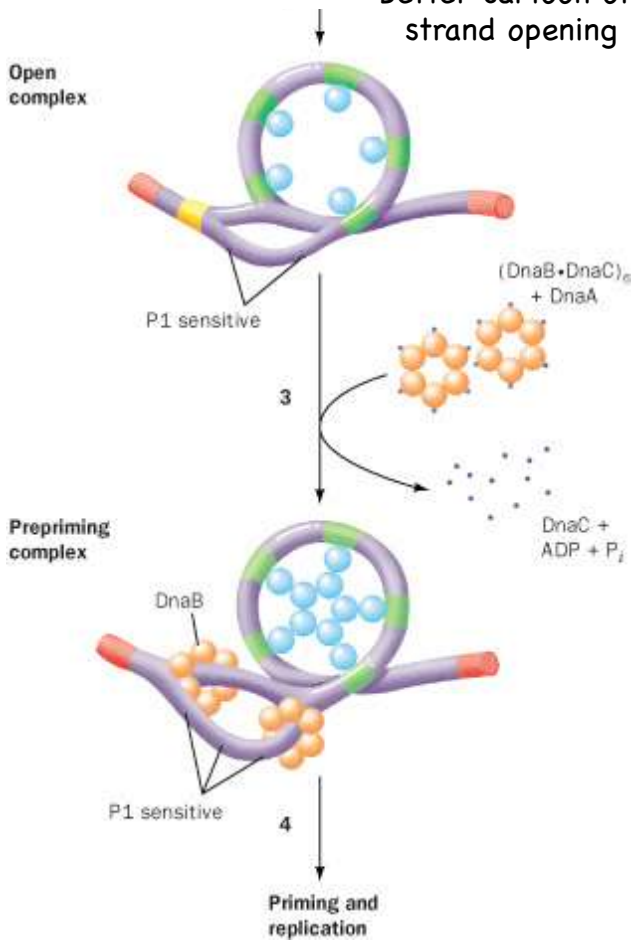



Hoatlin Fall 2009

DnaA binds to short repeated sequences and forms an oligomeric complex that melts DNA

6 DnaC monomers bind each hexamer of DnaB and this complex binds to the origin

Better cartoon of strand opening




DnaA is uniquely involved in initiation

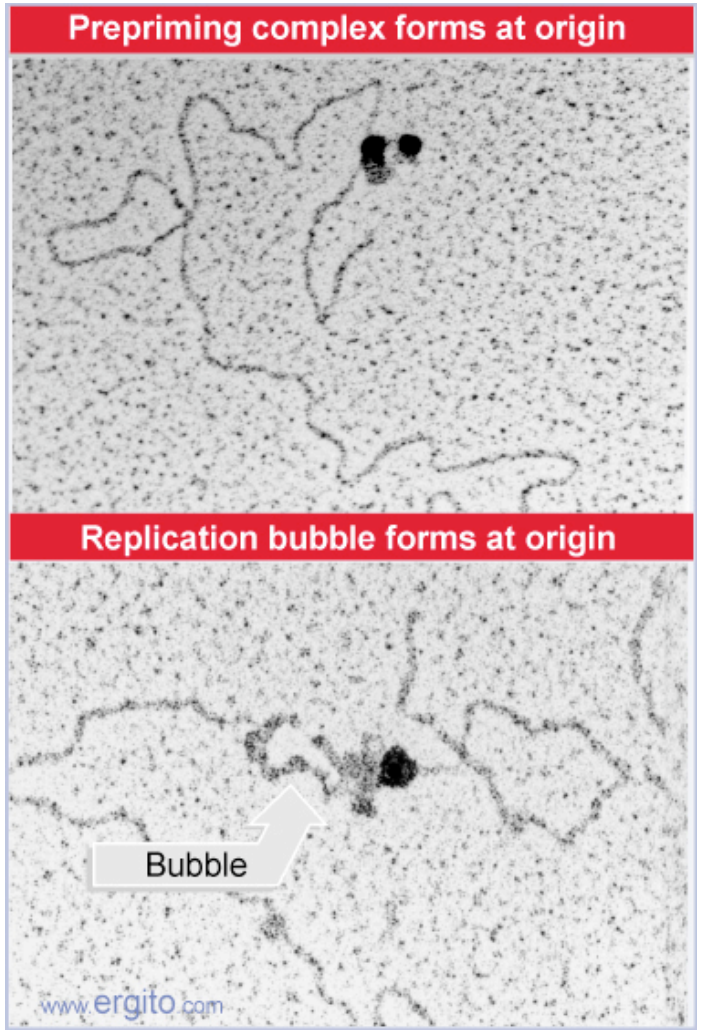
# Initiation

- DnaB initiates the two replication forks
- Displaces DnaA from the 13 bp repeats and extends the unwound region
- Each DnaB activates a DnaG primase
- DnaG primase initiates the leading strand and the first Okazaki fragment of the lagging strand

# Initiation, cont'd

- Gyrase-provides a swivel that allows one DNA strand to rotate around the other to reduce torsional strain of unwinding
- SSB- stabilizes the single strand DNA as it is formed
- About 60 bp needed to initiate replication

# Prepriming complex forms at origin




Pre-priming by sequential association of proteins.

The complex at *oriC* can be detected by electron microscopy with anti- DnaB

Hoatlin Fall 2009

Once started, replication progresses
for elongation bidirectionally
Forks meet
Control point only at initiation

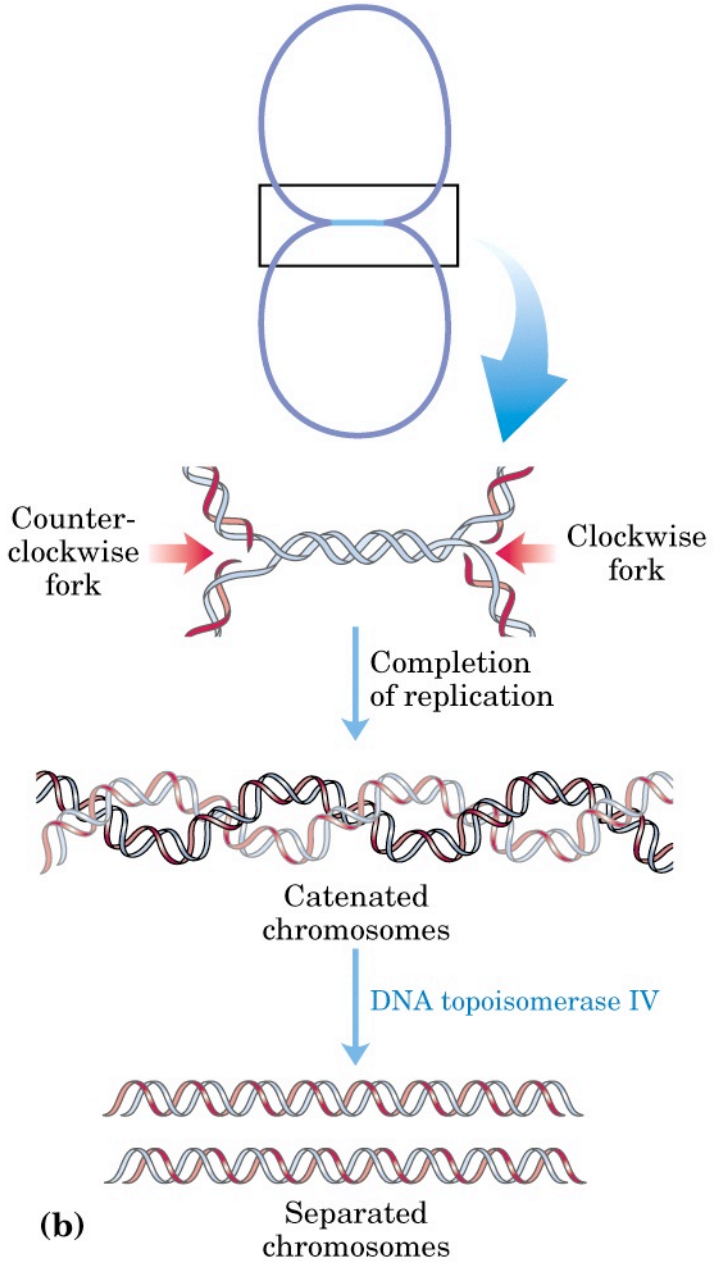


table 25–3

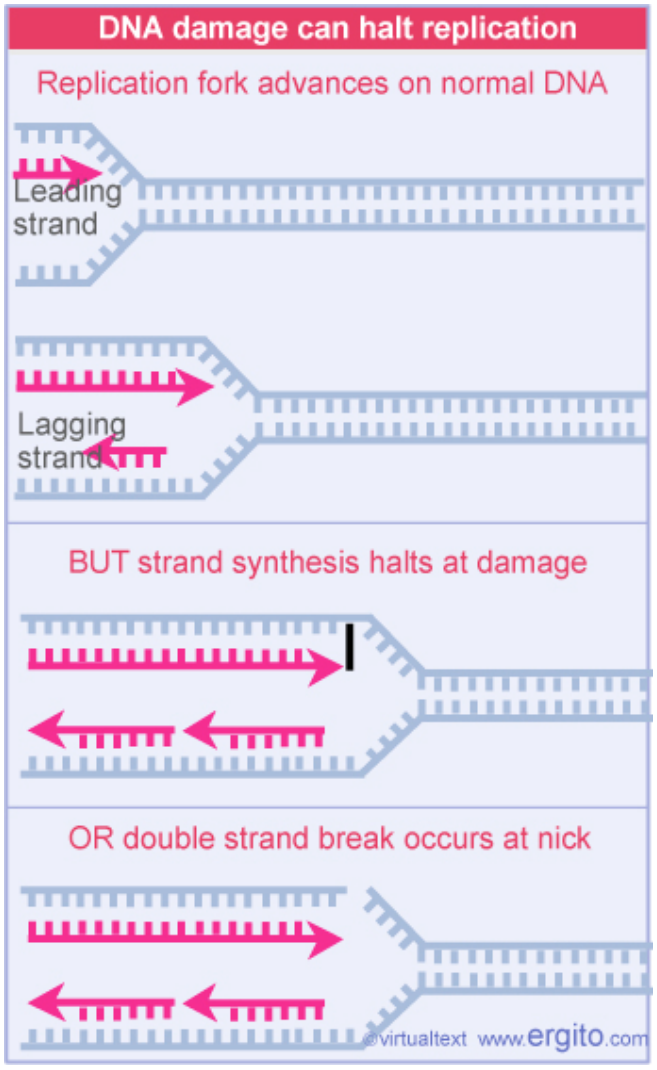
**Proteins Required to Initiate Replication at the *E. coli* Origin**

| Protein | $M_r$ | Number of subunits | Function |
|---|---|---|---|
| DnaA protein | 52,000 | 1 | Recognizes origin sequence; opens duplex at specific sites in origin |
| DnaB protein (helicase) | 300,000 | 6* | Unwinds DNA |
| DnaC protein | 29,000 | 1 | Required for DnaB binding at origin |
| HU | 19,000 | 2 | Histonelike protein; DNA bending protein; stimulates initiation |
| Primase (DnaG protein) | 60,000 | 1 | Synthesizes RNA primers |
| Single-stranded DNA-binding protein (SSB) | 75,600 | 4* | Binds single-stranded DNA |
| RNA polymerase | 454,000 | 5 | Facilitates DnaA activity |
| DNA gyrase (DNA topoisomerase II) | 400,000 | 4 | Relieves torsional strain generated by DNA unwinding |
| Dam methylase | 32,000 | 1 | Methylates (5′)GATC sequences at *ori*C |

*Subunits in these cases are identical.

- DnaA
  - Only used at initiation, ATP bound form is active
  - Mutations cause a slow-stop phenotype
  - Binds to set of 9 bp repeats
  - Further cooperative binding brings in 20 to 40 DnaA monomers
  - Melts the DNA at the 3- 13 bp repeats
- DnaB
  - ATP-dependent helicase
  - Displaces DnaA and unwinds DNA further to form replication forks
  - "Activates" primase, apparently by stabilizing a secondary structure in single-stranded DNA
- DnaC
  - Is in complex with DnaB before loading onto template
- Others: DnaG primase, Gyrase, SSB

# DNA damage can halt replication




When there is damage to a base in the DNA or a nick in one strand, DNA synthesis is halted, and the replication fork is either stalled or disrupted.

Stalling is very common.

It is not clear whether the components of the fork remain associated with the DNA or disassemble

# The primosome is needed to restart replication

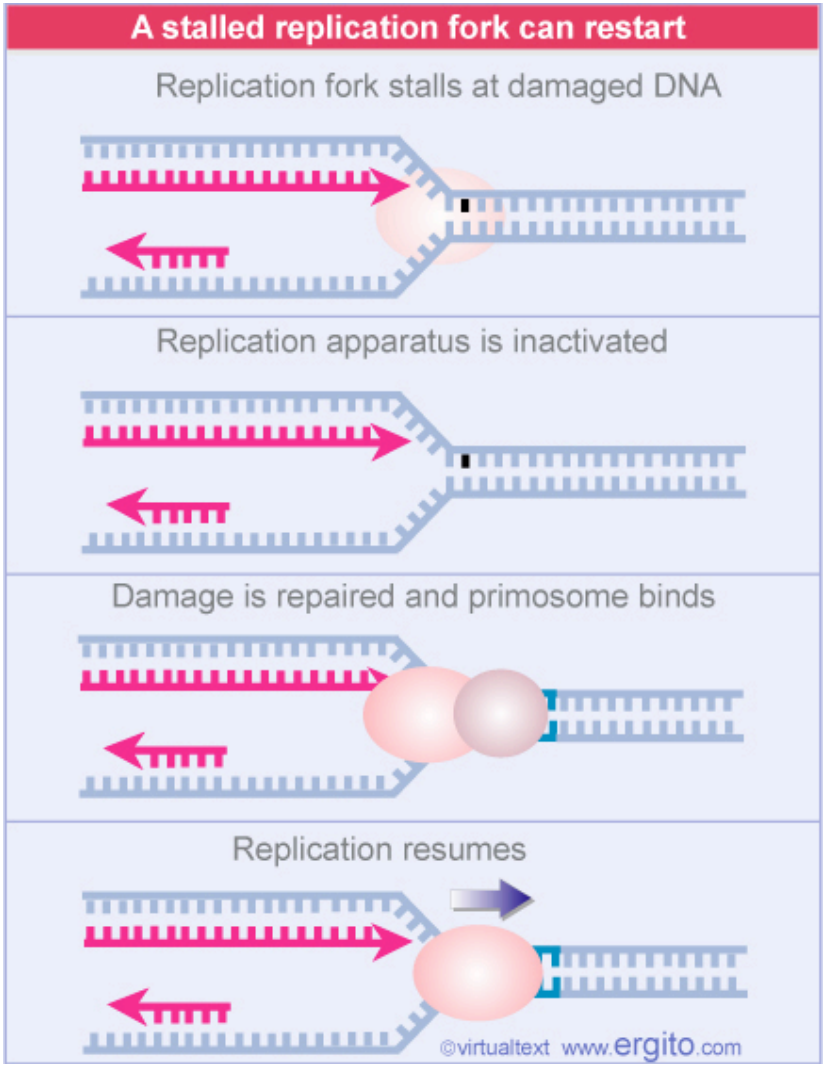
- The **primosome** describes the complex of proteins involved in the priming action that is involved in restarting stalled replication forks
- Primosome:PriA (3' to 5' helicase, site recog), PriB, PriC (adds DnaB-DnaC complex ), DnaT, DnaB (5' to 3' helicase, hexamer, DNA dependent ATPase), and DnaC (complexes with DnaB)
- PriA displaces SSB from ssDNA

New! Rewrite the textbooks! Heller& Marians. Mol Cell, 2005. PriC loads DnaB at a gap structure formed when a leading strand block occurs and the lagging strand continues, causing a gapped fork. Repriming the leading strand skips lesions, leaving gaps to be repaired later by recombination or trans-lesion polymerases

PriA loads replication proteins at another irregular structure to establish a new fork.

(lecture question....so what??)

# A stalled replication fork can restart




The primosome is required to restart a stalled replication fork after the DNA has been repaired. The primosome essentially reloads helicase DnaB in the absence of an origin so that helicase action can continue

Hoatlin Fall 2009

# Termination of replication: DNA sites and proteins needed

- DNA sites: *ter* sequences, 23 bp

– *terD* and *terA* block progress of counter-clockwise fork, allow clockwise fork to pass

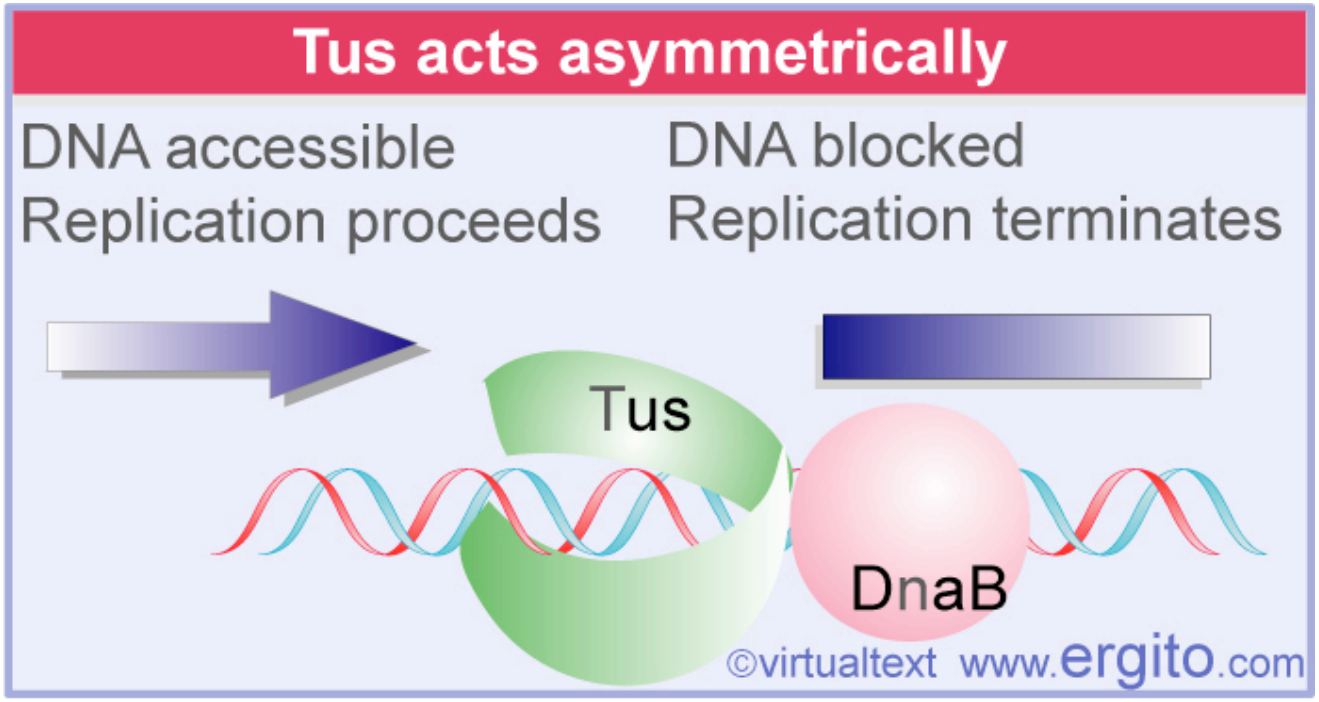
– *terC* and *terB* block progress of clockwise fork, allow counter- clockwise fork to pass

- Protein: Tus

– "*ter* utilization substance"

– Binds to *ter*

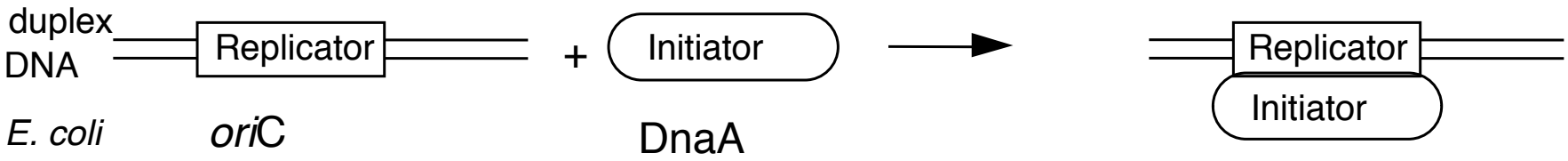
– Prevents helicase action from a specific replication fork




Tus binds to *ter* asymmetrically and blocks replication in only one direction

Ter and tus mutations are not lethal--not essential to terminate

# Replicon: a unit that controls replication




*E. coli* — *oriC* — DnaA

**Yeast** — ***ARS*** **an autonomously replicating sequence** — **ORC (the origin recognition complex) + ABF1 (ARS binding factor 1)** — **Complex of Replicator + initiator allows replication to begin**

Replicator: *cis*-acting DNA sequence required for initiation; defined genetically

Origin: site at which DNA replication initiates; defined biochemically

Initiator: protein needed for initiation, acts in *trans*

# Multiple replicons per chromosome

- Many replicons per chromosome, with many origins
- Replicons initiate at different times of S phase.
- Replicons containing actively transcribed genes replicate early, those with non-expressed genes replicate late.

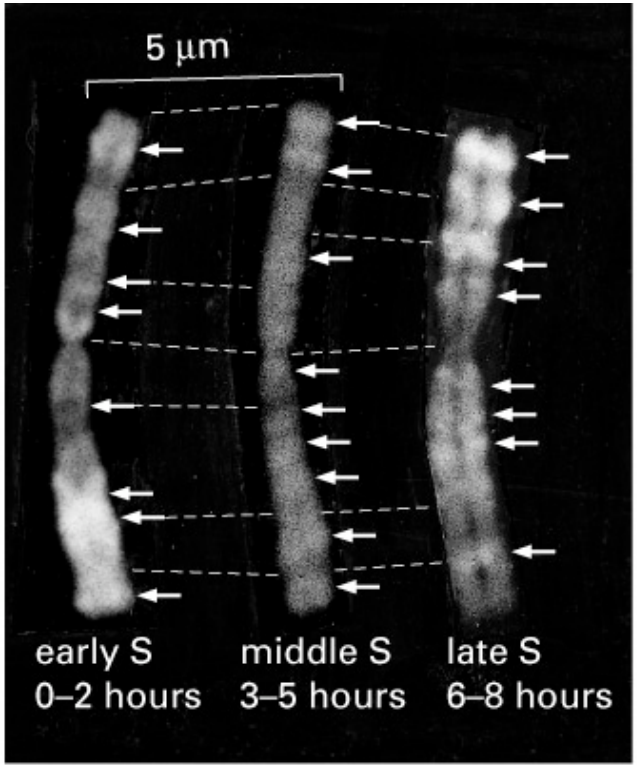
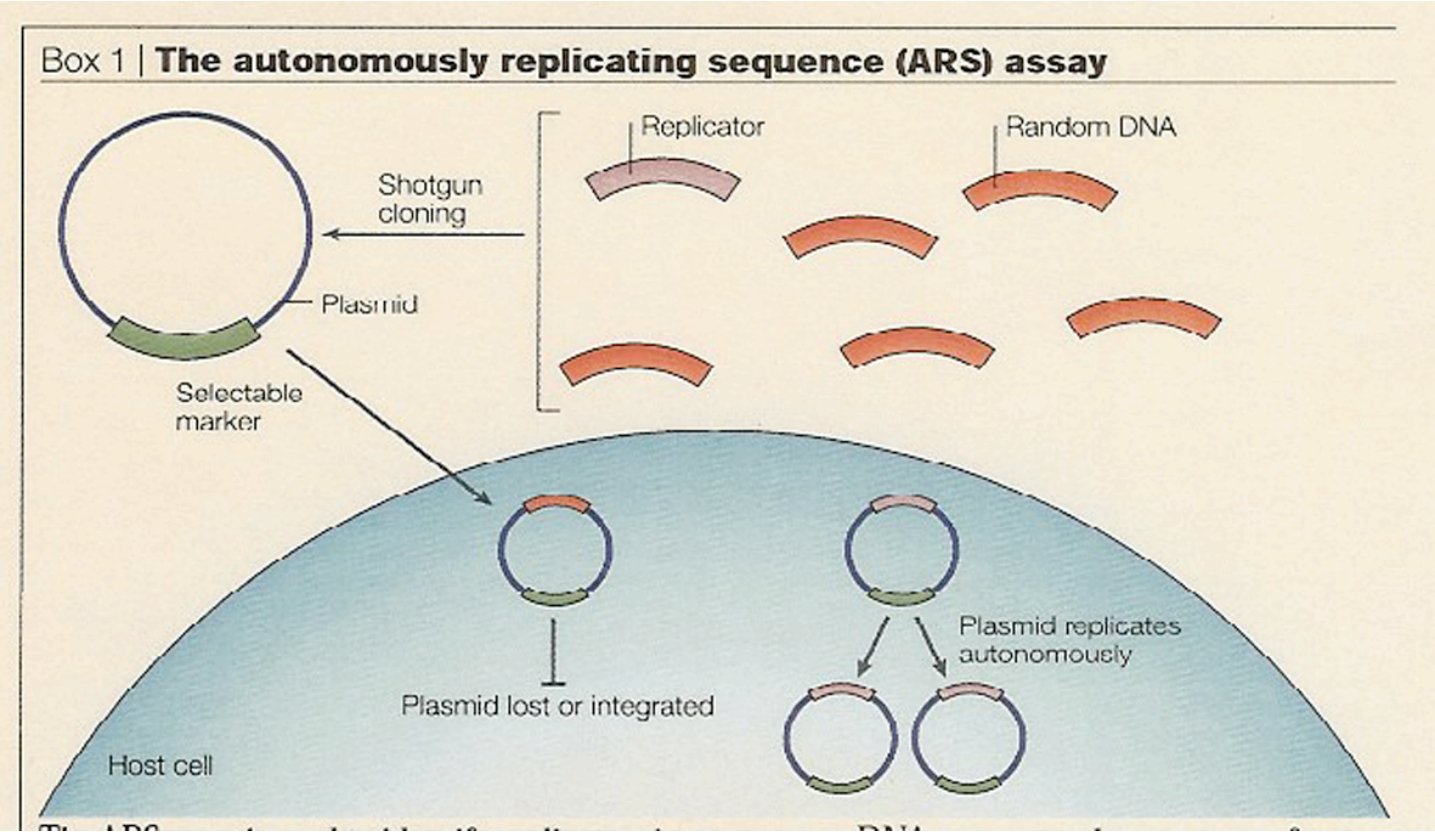



Figure 5–35. Molecular Biology of the Cell, 4th Edition.

# ARS Assay




"It is impossible to say how many clever permutations of this approach were tried and failed or found to be non-reproducible… the dirty 'random' word had reared its ugly head again, placing a nasty stigma on the ARS assay."

From, D. Gilbert, Nature Reviews Molecular Cell Biology, v5, 2004

# Acceptable Statements About Eukaryotic Replicons

**Replicators**

* There are many more replicators (that is, potential replication origins that are occupied by the origin-recognition complex (ORC)) than are used during S phase.
* More replicators are used than are needed to complete S phase.
* Replicators fire asynchronously throughout S phase.
* Replicator usage can be regulated during development and under different metabolic conditions.
* Once-per-cell-cycle genome duplication does not require specific replicator sequences.
* Different replicators have different DNA-sequence requirements; some have no DNA-sequence requirements.
* Many different features of chromatin structure and function can influence the activity of replicators.

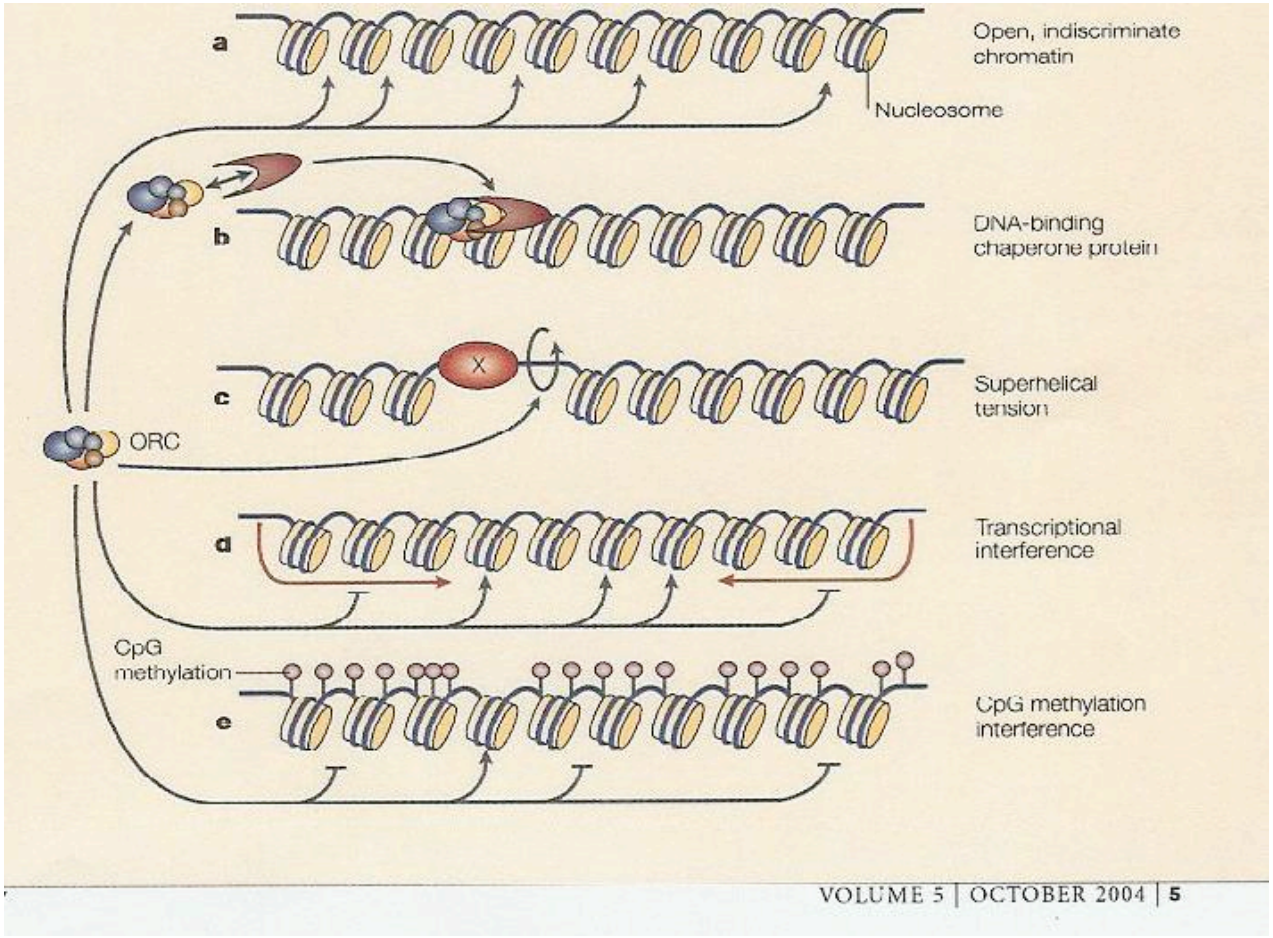
**Initiators**

* ORC homologues are found in all eukaryotes.
* ORC occupies replicators in vivo.
* The factors that influence the DNA-binding activity of ORC are highly divergent.

From, D. Gilbert, Nature Reviews Molecular Cell Biology, v5, 2004

Hoatlin Fall 2009

# Relaxed Replicon Model




Needs loading co-factor(s)

Geometry has to be OK

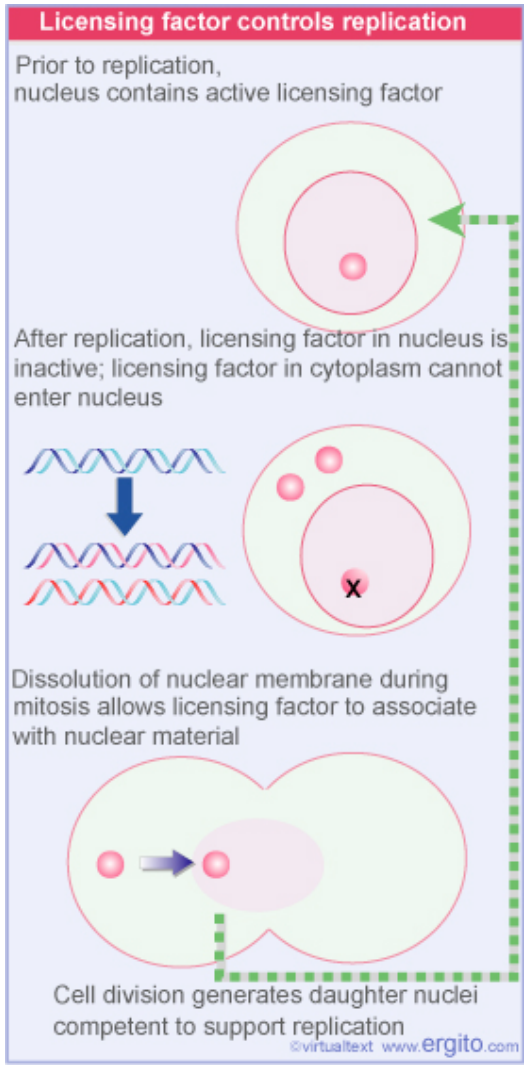
Transcription proteins may hinder access

Methylation may block

From, D. Gilbert, Nature Reviews Molecular Cell Biology, v5, 2004

# Licensing Factor Controls Eukaryotic Rereplication

- The origins in the multiple replicons fire once and only once in a single cell division
- **Licensing factor** is necessary for initiation of replication at each origins
- It is present in the nucleus prior to replication, but is inactivated or destroyed by replication
- Initiation of another replication cycle becomes possible only after licensing factor reenters the nucleus after mitosis

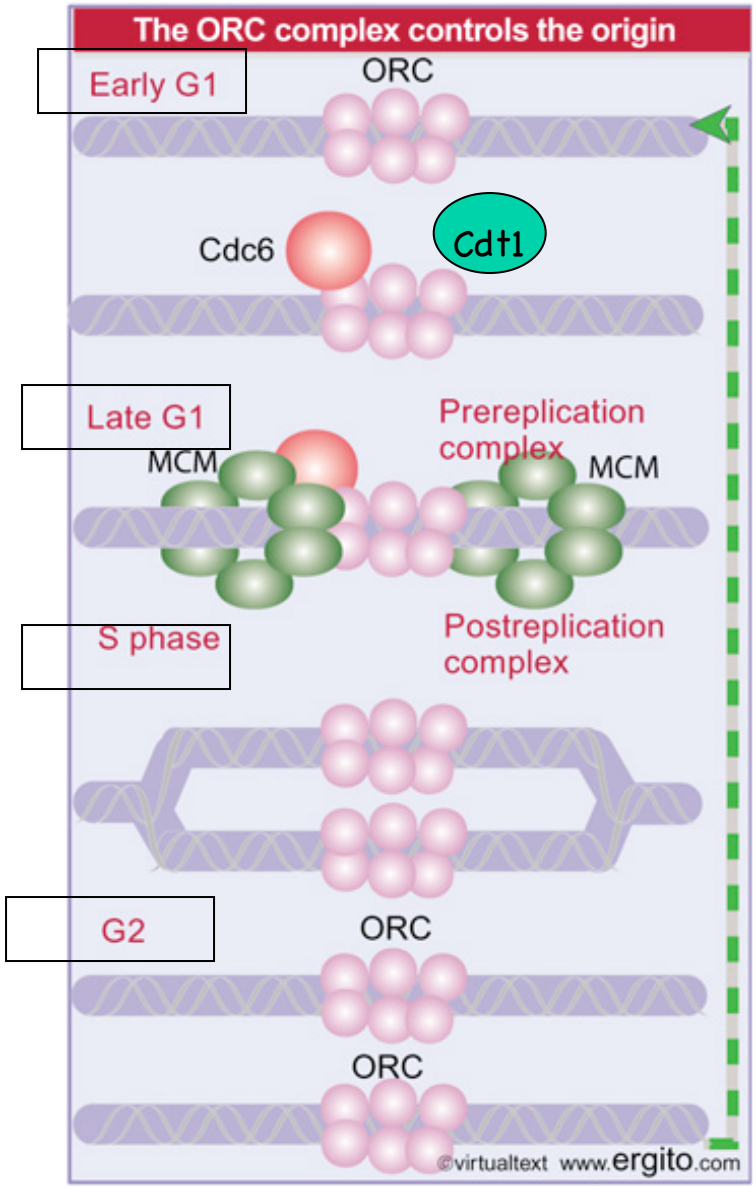



Licensing factor in the nucleus is inactivated after replication. A new supply of licensing factor can enter only when the nuclear membrane breaks down at mitosis

# Licensing Replication: The Players

- The ORC (origin recognition complex) is a 400 kD protein complex that is associated with yeast origins (ARS) throughout the cell cycle
- Changes occur in the pattern of DNA protected by ORC as a result protein-protein interactions with loading factors (e.g., Cdc6, Cdt1) and MCM
- The major role of ORC is to identify the origin to Cdc6 and MCM proteins that control initiation and licensing
- MCM 2-7 proteins form a 6 member ring-shaped complex MCM 4,6,7 have helicase activity, Mcm 2/3/7 are regulatory su.

# Proteins at the origin control susceptibility to initiation, Big Picture




Cdc6 is unstable, has ATPase activity, synthesized only in G1

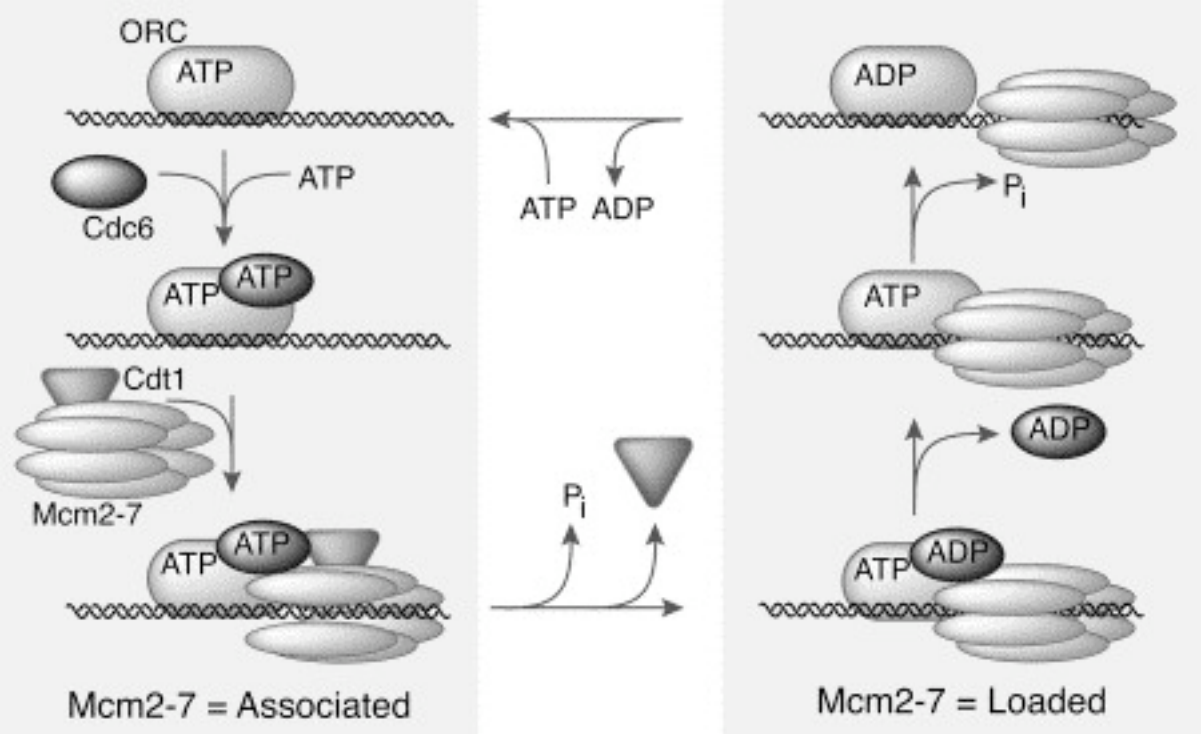
Loading factors Cdc6 & Cdt1 assembly on ORC, ORC position changes and allows MCM2-7 proteins to bind

When replication is initiated, loading factors are displaced, Cdc45 loads at each fork followed by phos Mcm➔ S phase

No re-initiation because loading factors are gone: degraded in yeast, phosphorylated and exported in mammalian cells

Hoatlin Fall 2009

# Proteins at the origin control susceptibility to initiation




1. ATP-bound ORC first binds origin DNA.
2. Cdc6 then binds ORC and ATP.
3. Cdt1 and Mcm2-7 (possibly as a complex) associate with ORC and Cdc6 at the origin.
4. ATP hydrolysis by Cdc6 leads to the loading of Mcm2-7 complexes on DNA and the release of Cdt1 from the origin.
5. Cdc6 association is destabilized by Cdc6 ATP hydrolysis.
6. ATP hydrolysis by ORC completes the Mcm2-7 loading reaction allowing further rounds of Mcm2-7 loading.

Randell at al 2006

# Introducing the MCM helicases

- MCM helicases forms double hexamer, modeled by SV40 large T antigen (Ltag)
- MCM s.u. zinc domain followed by AAA+ domain
- MCM looks like LTag in EM, but seq not similar
- Compare to the T7 hexamer ...NTP binding and hydrolysis "ripple" around ring leading to DNA translocation

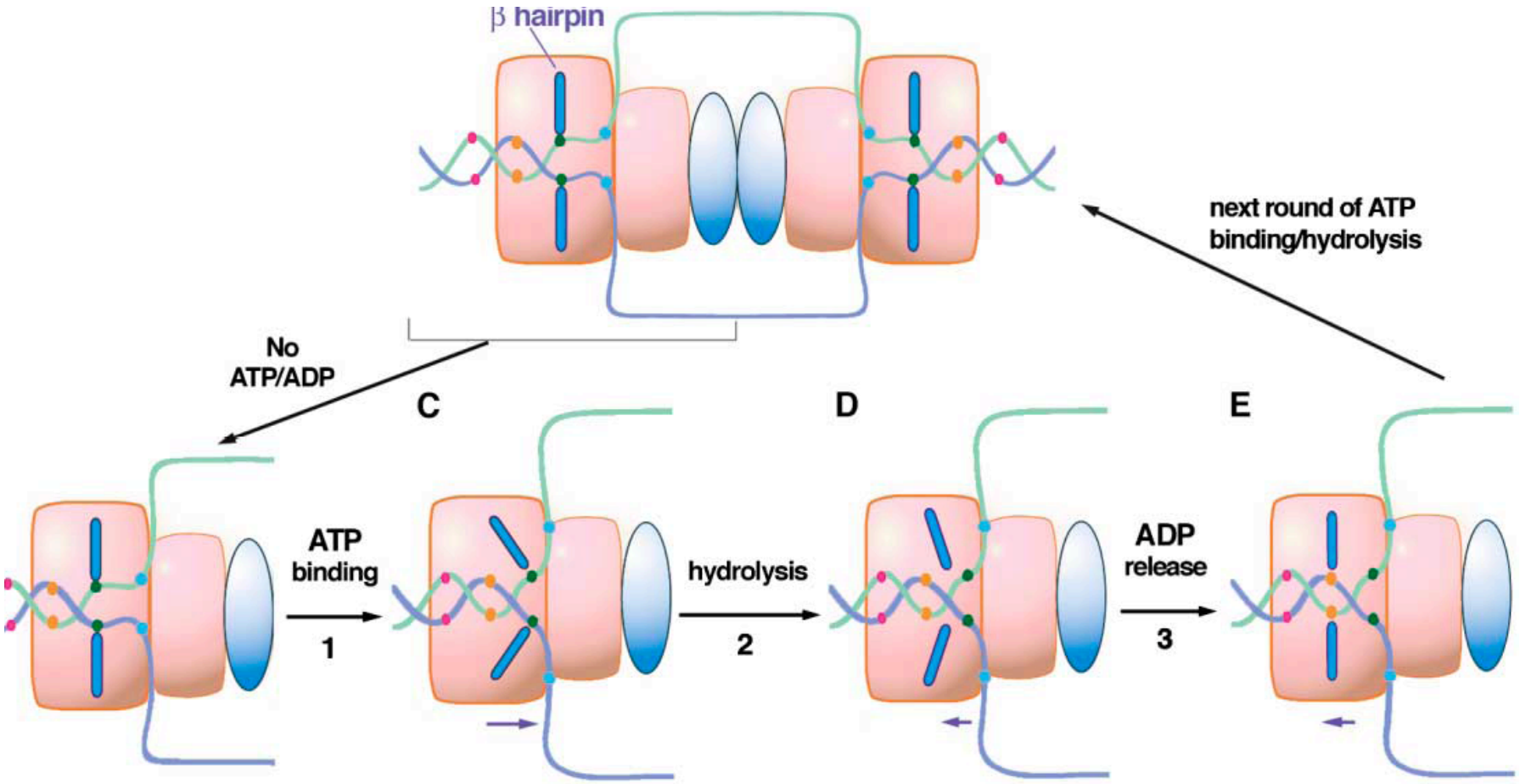
# Large T Ag Helicase

- LTag orchestrates
  - Recognition of SV40 viral origin of replication
  - assembles into a double hexamer
  - distorts origin locally
  - unwinds strands bidirectionally
  - recruits human initiation proteins

# Bidirectional Unwinding

- Duplex DNA pumped into double hexamer from both ends
- Two single stranded loops are extruded from the center
- Single stranded loops serve as templates for replication
- Mutations that prevent double hexamer prevent bidirectional unwinding and replication but not helicase activity of individual hexamers

# LTag Model




Gai, et al. Cell. 2004 Oct 1;119(1):47-60.

Hoatlin Fall 2009

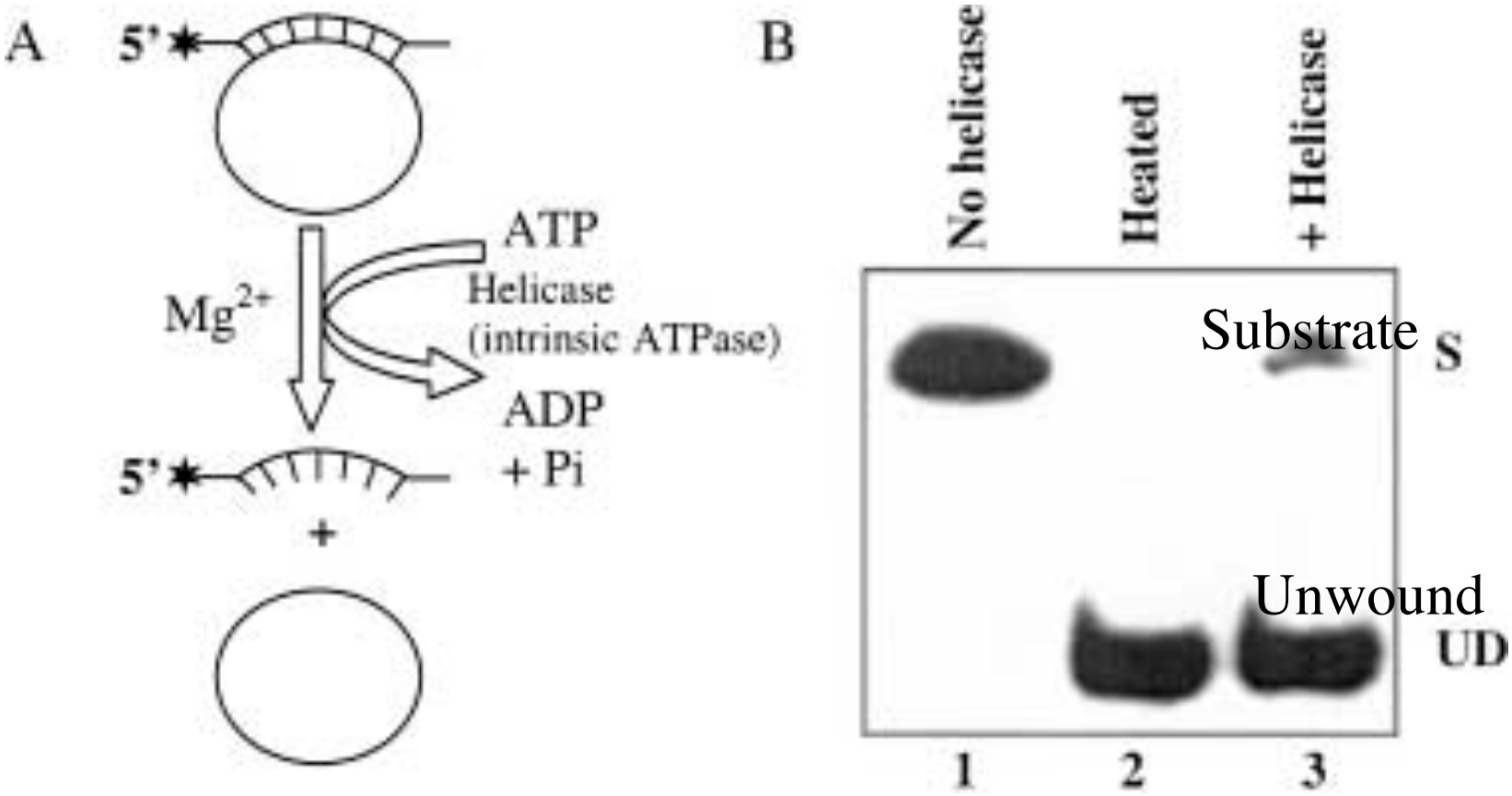
# The LTag (~MCM) Motor

- All 6 su bind ATP in concert
- Binding leads to conformational changes
- Iris-like motion narrows the central channel
- ATP hydrolysis causes rotational reversal
- Hypothesis: power stroke is iris opening and closing, driving translocation on DNA
- Basic β-hairpin loop in central channel binds origin DNA...loop moves 17 Å along channel when ATP binds

# Helicases

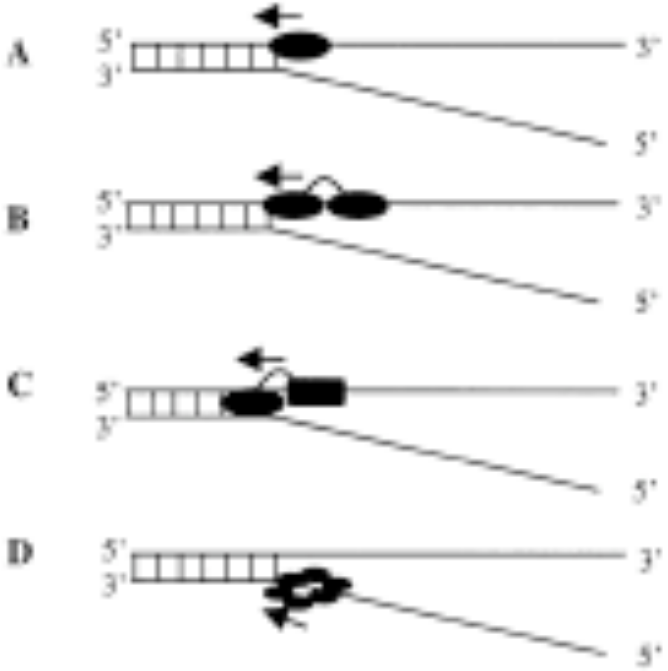
- Helicases are translocases and DNA-dependent ATPases
- Molecular motors using ATP hydrolysis to catalyze the unwinding of DNA duplex
- Roles in replication, repair, recombination, transcription
- Move unidirectionally along DNA

# General Unwinding Assay




Tuteja, Narendra & Tuteja, Renu (2004)
*European Journal of Biochemistry* **271** (10), 1835-1848.

# Forms of DNA Helicases at Forked Substrates




Bind both ss and ds DNA (PriA, UvrD)

1 su binds to ss DNA and motors

1 domain binds ds DNA and anchors other binds to ss and motors

Encircles DNA, binds ss at ss/ds junction (ex: BLM, SV40 LTAg, DnaB)

Tuteja, Narendra & Tuteja, Renu (2004)
*European Journal of Biochemistry* **271** (10), 1835-1848.

# Helicase Needs

- Usually require ssDNA loading zone-binding seq independent
- Exceptions are RecBCD and LTag, RuvB: all like ds DNA
- Many need a replication fork like structure
- Some (RecBCD,UvrD, Rep, RecQ, LTag) can unwind from blunt ended duplex DNA
- *E. coli* RecG, RuvA, and RuvB recognize Holliday junctions

# Human DNA Helicases

Table 8. Human DNA helicases. nd, Not determined.

| S. No. | Name of helicase | Mol. mass (kDa) | Nucleotide cofactors | Polarity | Remarks |
|---|---|---|---|---|---|
| 1. | HDH I[a] | 65 | ATP, dATP | 3′–5′ | Can also unwind DNA/RNA and RNA/RNA hybrids; may be involved in rDNA transcription. |
| 2. | HDH II/ Ku[b] | 87 72 | ATP, dATP | 3′–5′ | Functions in dsDNA break repair and V(D)J recombination; regulator of DNA-dependent protein kinase |
| 3. | HDH III[c] | 46 | ATP, dATP | 3′–5′ | Prefers replication fork like structure of substrates |
| 4. | HDH IV[d] | 100 | ATP, dATP | 5′–3′ | Can unwind DNA/RNA hybrids |
| 5. | HDH V[e] | 92 | ATP, dATP | 3′–5′ | Has highest turnover rate |
| 6. | HDH VI[f] | 128 | ATP, dATP | 3′–5′ | Prefers replication fork-like structure of substrates |
| 7. | HDH VII[g] | 36 | nd | nd | Trimer of one molecule of hnRNP A1 and two molecules of annexin II |
| 8. | HDH VIII[h] (G3GP) | 68 | ATP | 5′–3′ | A DNA and RNA helicase corresponding to G3 BP protein, an element of the RAS transduction pathway, similar to *E. coli* RecBCD |
| 9 | HDH IX[g] (RNP DO) | 45 | nd | nd | A Gly-Arg rich protein identified as ribonuclear protein DO |
| 10. | XPD/ERCC2[i] | 87 | ATP, dATP | 5′–3′ | Functions in nucleotide excision repair; component of BTF2-TFIIH transcription factor. |
| 11. | XPB/ ERCC3[j] | 89 | ATP | 3′–5′ | Functions in nucleotide excision repair; component of BTF2-TFIIH transcription factor. |
| 12. | Helicase ε[k] | 72 | ATP, dATP, CTP | 3′–5′ | Helicase activity is dependent on HRP-A |
| 13. | Helicase μ[l] | 110 90 | ATP, dATP, CTP, dCTP | 3′–5′ | Stimulated by 5′-tailed fork and SSB. |
| 14. | RIP 100[m] | 100 | ATP, dATP | 3′ 5′ | Associated with RIP60; RIP60 binds to replication origin region of DHFR[h] |
| 15. | Helicase Q1[n] | 73 | ATP, dATP | 3′–5′ | Gene homologous to *E. coli* RecQ gene; identical to human DNA helicase I |
| 16. | Helicase Q2[o] | 100 | ATP | 5′–3′ | Identical with DNA helicase IV |
| 17. | HehlR1[p] | 112 | ATP | 5′–3′ | Can unwind RNA/DNA substrates. Unlike others it can translocate along ssDNA in both directions when substrate have a very long ssDNA region. |
| 18. | HHcsA[q] | 116 | ATP, dATP | 5′–3′ | Hexameric protein. |
| 19 | WRN helicase[r] | 163 | ATP, dATP ≫ DCTP, CTP | 3′–5′ | Mutated in Werner syndrome, homologous to RecQ and contains 3′-5′ exonuclease activity |
| 20. | BLM helicase[s] | ≈ 160 | ATP | 3′–5′ | Mutated in cells of Bloom's syndrome patient and belongs to RecQ family. |
| 21. | Mcm4/6/7 complex[t] | ≈ 600 | ATP, dATP | 3′–5′ | The DNA unwinding activity is stimulated by SSB and forked DNA structures; can function as a replication helicase. |
| 22. | HEL308[u] | 124.5 | ATP, dATP | 3′–5′ | Homologous to DNA crosslink sensitivity protein Mus308 of *D. melanogaster*. Stimulated by RPA. |
| 23. | HFDH1[v] | ≈ 120 | ATP | 3′–5′ | First F-box protein that possesses enzyme activity. |
| 24. | Human RECQ1[w] | 75 | ATP | 3′–5′ | Needs 3′ tail of 10 nt on the substrate to open the duplex; can unwinds blunt end substrate with bubble of 25 nt; stimulated by hRPA |
| 25. | BACH1[x] | 130 | ATP | 5′–3′ | A nuclear phosphoprotein interacts with tumor suppressor, BRCA1. Involved in DSB repair and contain tumor suppression activity. |

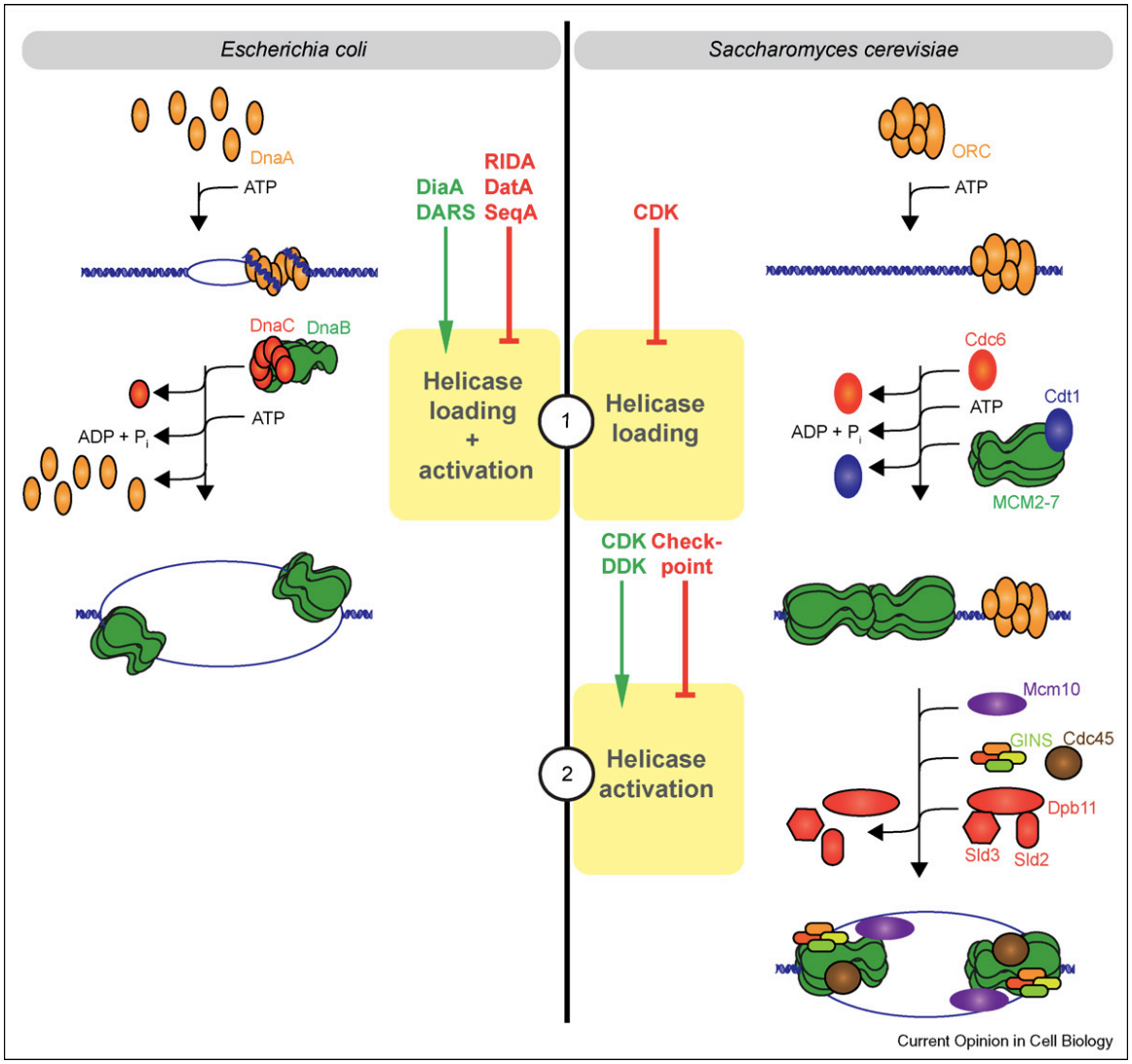
[a] [62]; [b] [112]; [c] [5]; [d] [5]; [e] [5]; [f] [5]; [g] A. Dallaire, B. A. Oelkers, unpublished data; [h] [114]; [i] [115]; [116]; [k] [117]; [l] [118]; [m] [119]; [n] [120]; [o] [120]; [p] [121]; [q] [122]; [r] [123]; [s] [124]; [t] [37]; [u] [125]; [v] [126]; [w] [51]; [x] [127].

Aka Brip1/FANCJ

Hoatlin Fall 2009

Tuteja, et al. FEBS 2004

# Stepwise loading and activation of the replicative helicase in Escherichia coli and Saccharomyces cerevisiae




Remus & Diffley 2009

Hoatlin Fall 2009

# Eukaryotic and Bacterial Comparisons

- DnaA and ORC (subunits 1–5) belong to the same clade of AAA+ ATPases
  - ORC subunits 1–5 and Cdc6 are structurally homologous to DnaA
- replicative helicases DnaB and Mcm2–7 are not orthologous
- DnaB and Mcm2–7 differ in their mode of DNA unwinding
- DnaB forms hexameric rings around the single-stranded lagging strand template of a replication fork and translocates in the 5′–3′ direction, sterically displacing the leading strand
- Mcm2–7: 3′–5′. similar to SV40 Ltag or bacterial recombination motor RuvB?
  - Requires Cdc45 and GINS
  - No DNA melting, ssDNA might not be required

Remus & Diffley 2009

# Eukaryotic and Bacterial Comparisons

- Loading onto single-stranded origin DNA is sufficient to activate E. coli DnaB
- Mcm2-7 is inactive after loading in M/G1 phase and requires activation in subsequent S phase
- Specific origin DNA sequences and origin melting do not appear to be required for helicase loading in eukaryotes
- Mechanism of CDK-dependent and DDK dependent activation of the Mcm2-7 helicase
  - induce the phosphorylation-dependent binding of Sld2 and Sld3 to Dpb11 (required for initiation but function unknown)
- DDK preferentially phosphorylates Mcm2-7 complexes that have been loaded onto DNA...then Cdc45 binds

Remus & Diffley 2009

# What remains to be discovered: prokaryotic

- how are 2 pols oriented, does helicase encircle one or 2 strands, is it a double hexamer, what happens when there is a pause/stall or a lesion, how is replication coordinated with repair and recombination? How are specialized pols delivered and unloaded?

# What remains to be discovered: eukaryotic

- how many more replisome factors are yet to be found? All replisomes the same or are there different types? How do DNA damage checkpoint mechanisms influence ongoing S-phase events? Are there orthologs of PriA and PriC that enable replication restart?....

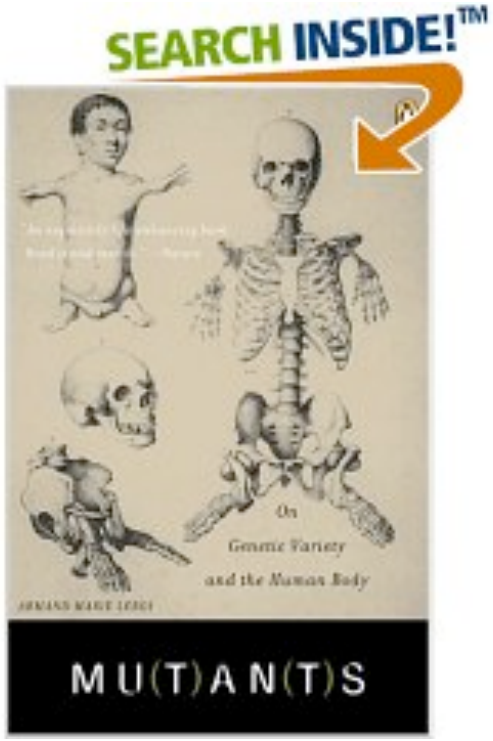
# Eukaryotic:Multiple Pathways Regulate Replication

- At least five protein pathways ensure that the genome is replicated *once and only once* during cell division
- Several additional pathways ensure that DNA replication is complete before mitosis begins

Adapted from DePamphilis 2006

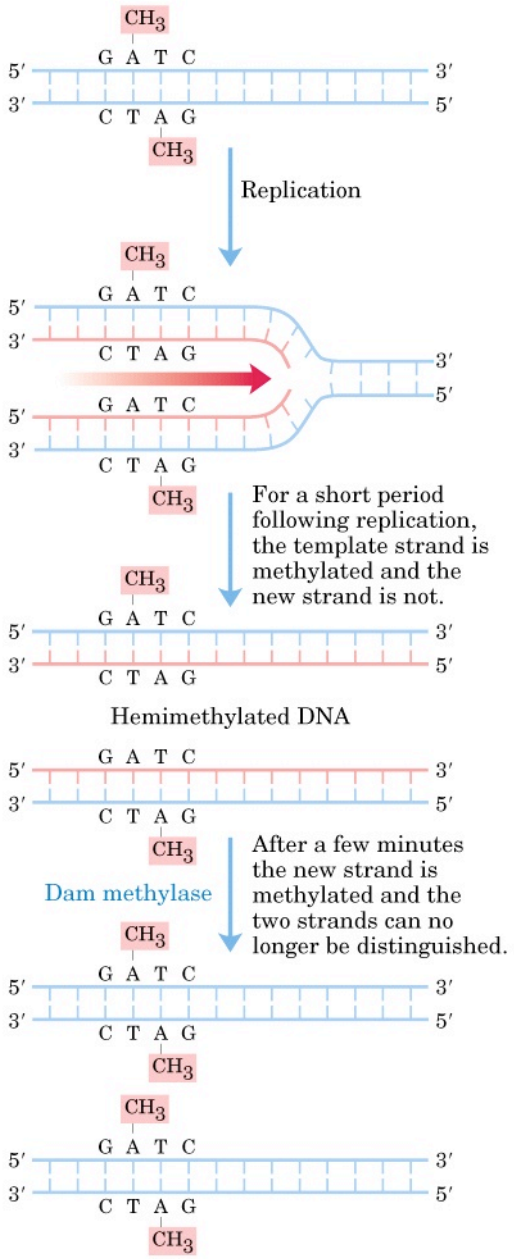
# Bonus Slides

# Great Book




**Mutants: On Genetic Variety and the Human Body by Armond Leroi, Available on Amazon, etc.**

Hoatlin CSF 2009

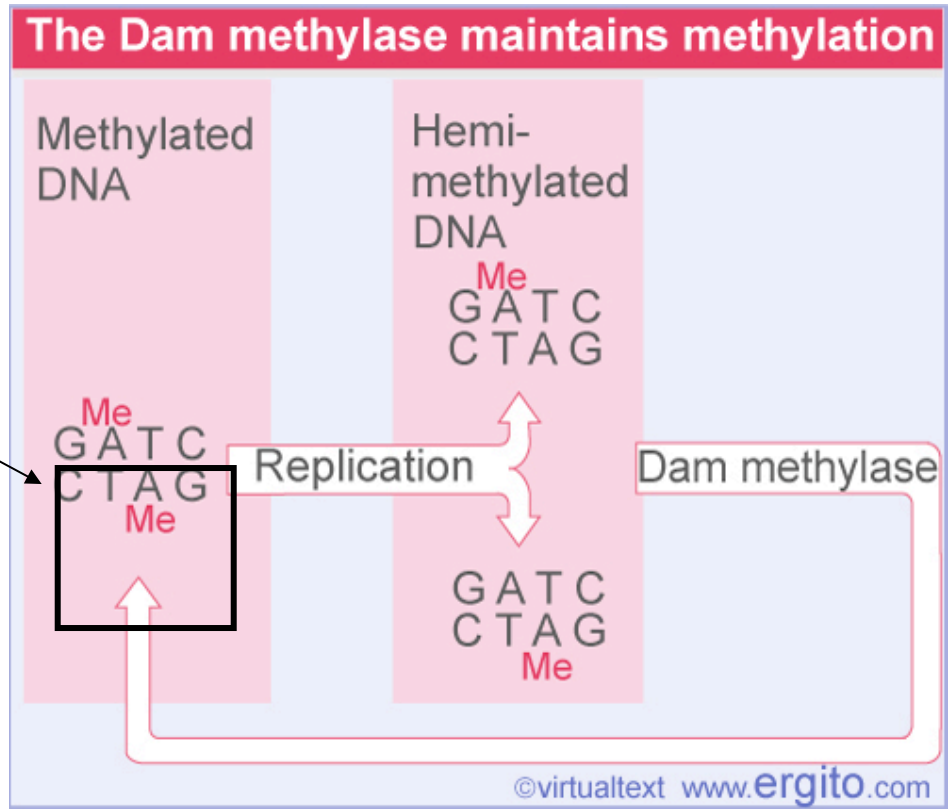



# Does methylation at the origin regulate initiation?

- *oriC* contains 11 [GATC] repeats that are methylated on adenine on both strands
- Replication generates hemimethylated DNA, which cannot initiate replication
- There is short delay before the repeats are remethylated

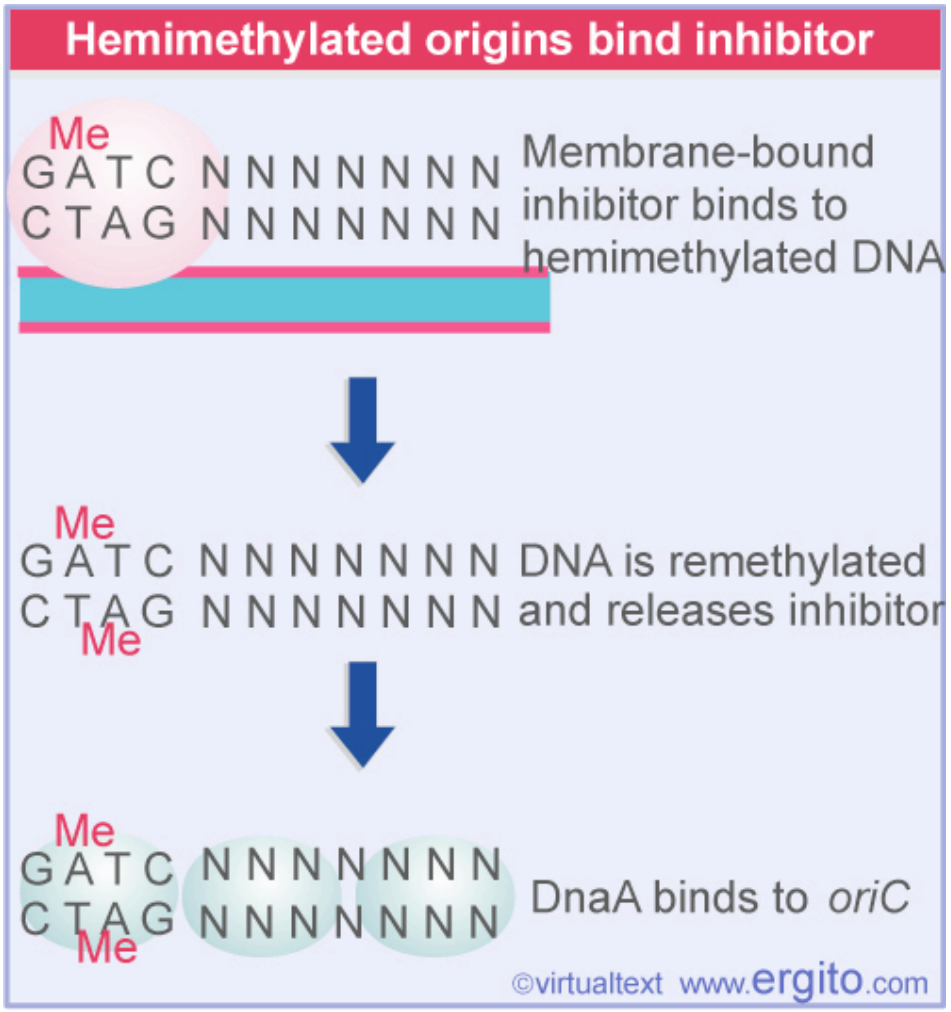
# Methylation Circuit

Target sequence for Dam Methylase in OriC repeats




Replication of methylated DNA gives hemimethylated DNA, which maintains its state at GATC sites until the Dam methylase restores the fully methylated condition

Hoatlin Fall 2009




A membrane-bound inhibitor binds to hemimethylated DNA at the origin, and may function by preventing the binding of DnaA. It is released when the DNA is remethylated

# Origins may be sequestered after replication

- While the origins are hemimethylated, they bind to the cell membrane, and may be unavailable to methylases
- The nature of the connection between the origin and the membrane is still unclear
- The only point at which *E. coli* can control DNA replication is at initiation--still not well understood